ABSTRACT

Title of Dissertation: TAMING THE SIGN PROBLEM IN LATTICE FIELD THEORY WITH DEFORMED PATH INTEGRAL CONTOURS

Neill Christian Warrington
Doctor of Philosophy in Physics
2019

Dissertation directed by: Professor Paulo Bedaque
Department of Physics

In this thesis a generic method for taming the sign problem is developed. The sign problem is the name given to the difficult task of numerically integrating a highly oscillatory integral, and the sign problem inhibits our ability to understand the properties of a wide range of systems of interest in theoretical physics. Particularly notably for nuclear physics, the sign problem prevents the calculation of the properties of QCD at finite baryon density, thereby precluding an understanding of the dense nuclear matter found in the center of a neutron star. The central idea developed in this thesis is to use the multidimensional generalization of Cauchy's Integral Theorem to deform the Feynman Path Integral of lattice fields theories into complexified field space to manifolds upon which the phase oscillations which cause the sign problem are gentle. Doing so allows calculations of theories with sign problems. Two practical manifold deformation methods, the holomorphic gradient flow and the sign-optimized manifold method, are developed. The holomorphic gradient flow, a generalization of the Lefschetz thimble method, continuously deforms the original path integration domain to a complex manifold via an evolution dictated by a complex first order differential equation. The sign-optimized manifold method is a way to generate a manifold with gentle phase oscillations by minimizing the sign problem in a parameterized family of manifolds through stochastic gradient ascent. With an eye toward QCD at finite density, the Cauchy's Theorem approach is applied to relativistic quantum field theories of fermions at finite density throughout this thesis. Finally, these methods are general and can be applied to both bosonic and fermionic theories, as well as Minkowski path integrals describing real-time dynamics.

TAMING THE SIGN PROBLEM
IN LATTICE FIELD THEORY
WITH DEFORMED PATH INTEGRAL CONTOURS

by

Neill C. Warrington

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Theoretical Physics
2019

Advisory Committee:

Professor Paulo Bedaque, Chair
Professor Tom Cohen
Professor Chirs Jarzynski
Professor Ian Appelbaum
Professor Don Perlis

Acknowledgements

I would like to thank a few people who have helped me on the way. I would like to thank my mother, who has been my biggest support my entire life. She loves nature more than anyone I've met, and my interest in science stems from her. I would like to thank my dad for always loving me, for being patient with me, and for pushing me in kendo and swimming. The lessons in hard work I learned through him make me who I am.

I would like to thank Kevin Croker for his friendship. He was the first person to show me the double slit experiment, he taught me relativity, and he's always believed in me. I would like to thank Tom Badman for his friendship. His art and his ideas in physics have shaped how I think, and he taught me how to draw and count music. I would like to thank Sara Fuentes for loving me and for supporting me in all that I do.

I would like to thank Tom Cohen for his constant support and for teaching me so much.

I would like to thank Andrei Alexandru. He was essentially my second advisor, he taught me so much and he pushed me to become a better physicist. He shaped how I think and he taught me nearly everything I know about computers.

Finally, I would like to Paulo Bedaque for being a truly wonderful advisor. He taught me everything I know, from dimensional analysis to quantum field theory. I'll never forget the day he taught me that 1 eV $\simeq 10^4$ K, and this is why the sun is red. He took me on when I knew nothing, he taught how to think, and

somehow extracted work out of me. Thank you so much Paulo. I have the chance to be a physicist because of you.

Table of Contents

List of Figures

Chapter 1

Introduction

The variety of phenomena that emerges from simple interactions between particles when large collections of particles are considered is striking [1]. From sound waves to superconductivity, the consequences of collective phenomena abounds.

The purpose of this thesis is to develop methods for quantum field theory which aid in the analysis of many-body systems. In particular, we develop methods for computing the properties of lattice quantum field theories at finite density. I will assume that the reader is familiar with the lattice formulation of quantum field theory throughout the text. The unfamiliar reader is referred to a few textbooks for an introduction [2–4].

To make the discussion concrete, suppose that one was interested in a particular many-body system, say a neutron star. A neutron star likely hosts very rich and interesting phenomena due to the extremely high densities achieved in its interior [5,6]. How might one calculate the properties of such a system from our most microscopic description of reality? To formulate the question precisely, we will approximate a neutron star as the equilibrium configuration of a large collection of baryons when all standard model processes other than quantum chromodynamics (QCD) are turned off. QCD is the quantum field theory describing the strong force, and this is not an unreasonable approximation given that a neutron star is composed primarily of baryons. Of course this is by no means the whole story, for without the weak interaction a neutron star could not even form[1], but the problem is difficult enough including only QCD at these high densities.

To compute the thermodynamic properties of a system governed by QCD with a non-zero density of baryons, one may utilize the formalism of (quantum) statistical mechanics [8]. It is a standard result of statistical mechanics [8,9] that all thermodynamic properties of a system can be computed from the partition function:

$$Z = \mathrm{tr}\left(\mathrm{e}^{-\beta(H-\mu Q)}\right) \ . \tag{1.1}$$

Here we are using the "grand-canonical ensemble", H is the Hamiltonian of QCD, β is the inverse temperature and Q is the baryon number operator

$$Q = \sum_f \int d^3x \ \bar{\psi}_f \gamma^0 \psi_f \ . \tag{1.2}$$

Here ψ_f are quark fields, f is a flavor index running over the six flavors of quarks, and μ is the chemical potential which is tuned to achieve the density of baryons desired. For those less familiar with this formalism, Q is an operator which counts the number of baryons minus the number of anti-baryons of a given state. For example, if $|\psi\rangle$ is the quantum state corresponding to a proton at rest, then $Q|\psi\rangle = (+1)|\psi\rangle$.

Though this is certainly a succinct procedure for computing the finite density thermodynamics of QCD, actually computing the partition function is rather daunting. How might one go about actually computing this partition function? One approach to this problem is to use lattice quantum field theory. It is a standard result [2–4,10] that the partition function of any field theory can be expressed

---

[1] There is substantial evidence that neutron stars are created when the cores of massive stars collapse due to gravitational forces. During the collapse, electrons and protons combine into neutrons and neutrinos through the weak interaction [7], and this process is responsible for the abundance of neutrons in the to-be neutron star.

as the continuum limit of a lattice path integral. That is,

$$Z = \mathrm{tr}\left(\mathrm{e}^{-\beta(H-\mu Q)}\right) \simeq \int D\bar{\psi} D\psi DA \ \mathrm{e}^{-S(\bar{\psi},\psi,A)} \tag{1.3}$$

where the equality becomes exact as the lattice spacing goes to zero. Notice the similarity between the expression Eq. (1.3) and the expression of the partition function in classical statistical mechanics of particles

$$Z = \int \prod_{i}^{N} d^3p_i d^3q_i \ \mathrm{e}^{-\beta H(p,q)} \ . \tag{1.4}$$

This formal similarity between the statistical mechanics of collections of classical particles and of quantum fields aids in the calculation of Eq. (1.3). In particular, numerical Monte Carlo methods developed to calculate observables from Eq. (1.4) can be also used to calculate observables from Eq. (1.3).

We have carried out two steps in writing the partition function as a path integral. First, the continuum action is "latticeized". To carry out this procedure, one takes the continuum action and replaces the integral over spacetime with a finite sum over a hypercubic lattice of finite volume and non-zero lattice spacing a. That is

$$\int d^4x L(\bar{\psi}, \psi, A) \rightarrow \sum_x a^4 L(\bar{\psi}_x, \psi_x, A_x) \tag{1.5}$$

where $\bar{\psi}_x, \psi_x$ and A_x are fermion fields and gauge fields on lattice site x respectively. The lattice serves as a UV regulator and restricts energies and momenta to be less than the inverse lattice spacing, thus rendering observables finite and free of divergences. Continuum physics is recovered by taking the lattice spacing to zero. The technical details of taking the continuum limit are non-trivial and we will not discuss them here. The interested reader may consult the standard texts [2–4, 10]. Once the theory is put on the lattice, the partition function can be written as a path integral by the repeated insertion of coherent states of the gauge and fermion fields. This "coherent state path integral" is a standard construction [3].

Having put the theory on the lattice, and having expressed the lattice partition function as a (coherent state) path integral, we are left with the expression

$$Z = \int D\bar{\psi} D\psi DA \; \mathrm{e}^{-S(\bar{\psi},\psi,A)} \; . \tag{1.6}$$

Here, S is the Euclidean action which takes the form

$$S(\bar{\psi}, \psi, A) = S_g(A) + \sum_{xy} \bar{\psi}_x D_{xy}(A) \psi_y \tag{1.7}$$

where S_g is the "gauge action" which encodes the dynamics of the bosonic fields A, and where D is the fermion matrix, which governs the dynamics of the fermions coupled to the bosonic fields. Throughout this thesis we will use the shorthand

$$D\bar{\psi} D\psi \equiv \prod_x d\bar{\psi}_x d\psi_x, \quad DA \equiv \prod_x dA_x \tag{1.8}$$

for brevity. Here $\prod_x$ ranges over all lattice sites. The fermionic integration measure $D\bar{\psi} D\psi$ is over grassmann variables [3] at each lattice site and the bosonic integration measure DA is over gauge fields on all links. Integrating over the fermions yields a determinant of the fermion matrix [3] and one has

$$Z = \int DA \; \mathrm{e}^{-S_g(A)} \mathrm{det} D(A) \; . \tag{1.9}$$

When calculating observables from Z numerically, the expression above, which involves only real numbers (rather than Grassmann variables), is what is evaluated. From this partition function, any observable in thermal equilibrium can be evaluated in the grand canonical ensemble

$$\langle \mathcal{O} \rangle = \frac{\int DA\ \mathcal{O}(A)\ \mathrm{e}^{-S_g(A)} \mathrm{det} D(A)}{\int DA\ \mathrm{e}^{-S_g(A)} \mathrm{det} D(A)} \ . \tag{1.10}$$

One can calculate, for example, the density $\langle n \rangle = \langle Q \rangle / V$ of baryon number, or the compressibility $\kappa = \frac{\partial}{\partial V}\left(\frac{-T \ln Z}{V}\right)$ of this material, as a function of temperature and chemical potential. Knowing these functions are useful in characterizing the properties of this material and determine, for instance, the mass-radius relationship for a neutron star [11].

How does one compute these path integrals in practice? Exact evaluation of the high dimensional integrals we come across in lattice field theory is almost always prohibitively difficult. However, in the $\mu = 0$ case, there is a well-developed art for computing lattice path integrals numerically with Monte Carlo methods [2–4, 10].

To evaluate Eq. (1.10), one approach is to first notice that

$$p(A) = \frac{e^{-S_g(A)} \mathrm{det} D(A)}{Z} \equiv \frac{e^{-S(A)}}{Z} \tag{1.11}$$

can be interpreted as a probability distribution provided $\mathrm{det} D(A) \geq 0$ for any A. This is typically the case for $\mu = 0$. Note that we have defined the notation $S(A) = S_g(A) - \ln\, \mathrm{det} D(A)$ for convenience. Therefore we have

$$\langle \mathcal{O} \rangle = \int DA\ p(A) \mathcal{O}(A) \ , \tag{1.12}$$

and so if we collect a number of field configurations $\{A_1, ..., A_{\mathcal{N}}\}$ distributed according to the probability distribution $p(A)$, then we can estimate observables as:

$$\langle \mathcal{O} \rangle \simeq \frac{1}{\mathcal{N}} \sum_i \mathcal{O}(A_i) \ , \tag{1.13}$$

with a statistical uncertainty in $\langle \mathcal{O} \rangle$ proportional to $\frac{1}{\sqrt{\mathcal{N}}}$. There are many methods for sampling a probability distribution. In this thesis we use a family of methods called "Markov Chain Monte Carlo" (MCMC). The reader interested in the details of how to carry out an MCMC is directed to the review [12].

What about $\mu \neq 0$? Then a complication arises: the fermion determinant becomes complex. In this case, $\frac{e^{-S(A)}}{Z}$ cannot be interpreted as a probability distribution, and so the Monte Carlo method described above is not valid. Theories with complex actions are said to have a "sign problem". Theories with sign problems can be found across various fields of physics from QCD at finite density [13,14] to the Hubbard model away from half filling [15].

So what can be done? There are many approaches to dealing with the sign problem. A common approach is to define

$$p(A) = \frac{e^{-\text{Re } S(A)}}{\int DA\ e^{-\text{Re } S(A)}} \ , \tag{1.14}$$

which *is* a probability distribution, and collect field configurations according to $\frac{e^{-\text{Re } S(A)}}{\int DA\ e^{-\text{Re } S(A)}}$, while absorbing the phase $\mathrm{e}^{-i\text{Im } S}$ into observables. This procedure is called "reweighting" [16]. To be explicit, reweighting amounts to the following exact decomposition:

$$\begin{aligned}\langle \mathcal{O} \rangle &= \frac{\int DA\ \mathrm{e}^{-S(A)}\mathcal{O}(A)}{\int DA\ \mathrm{e}^{-S(A)}} = \frac{\int DA\ \mathrm{e}^{-\text{Re } S(A)}\mathrm{e}^{-i\text{Im } S(A)}\mathcal{O}(A)}{\int DA\ \mathrm{e}^{-\text{Re } S(A)}\mathrm{e}^{-i\text{Im } S(A)}} \\ &= \frac{\int DA\ \mathrm{e}^{-\text{Re } S(A)}\mathrm{e}^{-i\text{Im } S(A)}\mathcal{O}(A) / \int DA\ \mathrm{e}^{-\text{Re } S(A)}}{\int DA\ \mathrm{e}^{-\text{Re } S(A)}\mathrm{e}^{-i\text{Im } S(A)} / \int DA\ \mathrm{e}^{-\text{Re } S(A)}} \\ &= \frac{\langle \mathcal{O}\mathrm{e}^{-i\text{Im } S}\rangle_{\text{Re } S}}{\langle \mathrm{e}^{-i\text{Im } S}\rangle_{\text{Re } S}}\end{aligned} \tag{1.15}$$

where $\langle \ \cdot \ \rangle_{\text{Re } S}$ mean to average with respect to the distribution Eq. (1.14). The name "reweighting" makes sense: any observable can be expressed as a ratio of observables averaged with respect to the real part of the action, reweighted by the fluctuating phase $\mathrm{e}^{-i\mathrm{Im}\ S}$.

This general procedure of reweighting allows one to compute the properties of theories with complex actions. There's just one catch: the denominator $\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}$ has a magnitude that gets small exponentially fast in both the spatial volume and the inverse temperature. Moreover, the standard deviation of $|\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}|$ remains order one (because it an average of numbers which live on the unit circle) as the mean decreases [17]. This combination of effects, the variance of the average sign being much larger than an exponentially small mean is what makes theories with sign problems difficult to analyze.

One approach to demonstrating the exponential decay of the average sign is given in [17], and I will reproduce it here. The average phase is a ratio of partition functions:

$$\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S} = \frac{\int DA\ \mathrm{e}^{-S}}{\int DA\ \mathrm{e}^{-\mathrm{Re}\ S}} = \frac{Z}{Z'} \tag{1.16}$$

Using the relation $Z = \mathrm{e}^{-\beta F}$ where F is a free energy, we then find

$$\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S} = \mathrm{e}^{-\beta(F-F')}\ . \tag{1.17}$$

where F is only guaranteed to real in the continuum limit[2]. Note however that $|F| > |F'|$ in general because

$$|Z| = |\int DA\ \mathrm{e}^{-S}| = |\int DA\ \mathrm{e}^{-S}| \leq \int DA\ |\mathrm{e}^{-S}| = \int DA\ \mathrm{e}^{-\mathrm{Re}\ S} = Z' \tag{1.18}$$

Now the free energy, being an extensive quantity, can be written as $F = Vf$ where f is finite in the thermodynamic limit. Therefore we find that

$$|\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}| = \mathrm{e}^{-\beta V(f-f')}\ , \tag{1.19}$$

---

[2] In the continuum limit $Z = \int DA\ \mathrm{e}^{-S} \to \mathrm{tr}\big(\mathrm{e}^{-\beta(H-\mu N)}\big)$ which is manifestly real because it is the sum of eigenvalues all of which are real and positive.

showing that the modulus of the average phase vanishes exponentially fast in βV.

The exponential smallness of the average phase makes it very difficult to estimate observables computed with a Monte Carlo because the statistical uncertainty in a stochastically estimated quantity scales as $\mathcal{N}^{-1/2}$ [12]. This means that $\mathcal{O}(\mathrm{e}^{\beta V(f-f')^2})$ samples are required to estimate the denominator of Eq. (1.15) [17]. This has the consequence that, until an exponential number of field configurations are collected, one basically has no idea what the value of an observable is. This is a simple point, but it is worthwhile to dwell on this for a moment. Suppose we are calculating observables in a theory with a sign problem and sufficiently many configurations are collected to determine that $\mathrm{Re}\ \langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S} = 0.001 \pm 0.1$. Then

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O}\mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}}{0.001 \pm 0.1}\ . \tag{1.20}$$

This means that the observable is completely unknown since the denominator overlaps with zero. Obviously this is not a useful result, and this demonstrates why theories with a sign problem are so difficult to analyze: until $\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}$, a quantity exponentially small in the spacetime volume, is resolved from zero, *all* observables are unknown.

Because the quantity $\langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}$ determines the severity of the sign problem, it will come up frequently in the text to follow. Therefore I will give it the name "the average phase" or "the average sign" and we define the shorthand $\langle \sigma \rangle \equiv \langle \mathrm{e}^{-i\mathrm{Im}\ S}\rangle_{\mathrm{Re}\ S}$.

This brings us to the main idea of the present thesis: I will explain a generic method for taming the sign problem [18]. By this I mean that I will describe a procedure which can be applied to a large number of theories that takes the average sign and makes it larger [19–23]. Since the average sign is a quantity which controls the severity of the sign problem, when the average sign is increased a theory becomes “easier” to solve. Our method sometimes improves the average sign dramatically, and renders previously unsolvable theories solvable.

To be clear, I do not have a solution to the sign problem. Indeed it is known that the sign problem is NP-hard [17], implying that there is no generic solution unless $NP = P$. Instead, what I will present is a method for *alleviating* the sign problem. By alleviate I mean the following. Recall that the average phase can be expressed as

$$\langle \mathrm{e}^{-i\mathrm{Im}\ S} \rangle_{\mathrm{Re}\ S} = \mathrm{e}^{-\beta V(f-f')} \tag{1.21}$$

In essence, the method I will describe is a way to take the free energy difference $f - f'$ and make it smaller. It is in this way that the sign problem is alleviated. Sometimes the free energy difference can be shrunk so much that a problem which was, practically speaking, unsolvable, becomes solvable. Of course at large enough volumes, the average phase will always vanish exponentially, but infinite volumes aren’t necessary in lattice calculations. All that are needed are volumes large enough that an extrapolation can be performed.

I will briefly explain the method here and the details are in the chapters to follow. The central idea at the heart of the method is the following theorem.

Suppose $f = f(x)$ is an analytic function of $x \in \mathbb{R}^N$. Then f has an analytic continuation $f = f(z)$ for $z \in \mathbb{C}^N$. Such a function $f(z)$ has the property $\frac{\partial f}{\partial \bar{z}_a} = 0$, where $\bar{z}_a$ is the complex conjugate of z_a, for every a. Then

$$\int_{\mathbb{R}^N} f(x)\ dx_1 \wedge dx_2 \wedge ... \wedge dx_N = \int_{\mathcal{M}_0 \subset \mathbb{C}^N} f(z)\ dz_1 \wedge dz_2 \wedge ... \wedge dz_N\ , \tag{1.22}$$

where $M_0 = \mathbb{R}^N$. We are now viewing the integral as being performed over $\mathbb{R}^N$, a submanifold of $\mathbb{C}^N$. The manifold $\mathbb{R}^N$ is real dimension N while $\mathbb{C}^N$ is real dimension $2N$. Consider smoothly deforming $\mathcal{M}_0$ into $\mathcal{M}_T \subset \mathbb{C}^N$. At each step the deformed manifold will be a complex manifold with real dimension N, therefore the sequence of deformations sweeps out a $N+1$ dimensional manifold which we will call Σ. Then the boundary of Σ is $\partial\Sigma = \mathcal{M}_0 \cup (-\mathcal{M}_T)$. Stokes' theorem implies

$$\begin{aligned}\int_{\partial\Sigma} f(z)\; dz_1 \wedge dz_2 \wedge ... \wedge dz_N &= \int_{\Sigma} \sum_a \big[\frac{\partial f}{\partial z_a} dz_a + \underbrace{\frac{\partial f}{\partial \bar{z}_a} d\bar{z}_a}_{=0}\big] \wedge dz_1 \wedge dz_2 \wedge ... \wedge dz_N \\ &= \sum_a \int_{\Sigma} \frac{\partial f}{\partial z_a} dz_a \wedge dz_1 \wedge dz_2 \wedge ... \wedge dz_N \\ &= 0\ .\end{aligned} \tag{1.23}$$

Therefore

$$\int_{\mathcal{M}_0} f(z)\; dz_1 \wedge dz_2 \wedge ... \wedge dz_N = \int_{\mathcal{M}_T} f(z)\; dz_1 \wedge dz_2 \wedge ... \wedge dz_N\ . \tag{1.24}$$

What this means is that integrals of holomorphic functions over $\mathbb{R}^N$ can be deformed to integrals over manifolds in $\mathbb{C}^N$. We will call this "Cauchy's theorem in many dimensions" or just "Cauchy's theorem".

The idea is, since lattice path integrals are just many dimensional integrals of (often) holomorphic functions, Cauchy's theorem applies[3]. The partition function can be computed over a host of manifolds in $\mathbb{C}^N$:

$$Z = \int_{\mathbb{R}^N} DA\; e^{-S} = \int_{\mathcal{M}\subset\mathbb{C}^N} DA\; e^{-S} \tag{1.25}$$

A corollary of Cauchy's theorem is that holomorphic observables can be integrated over deformed contours as well. That is, if $\mathcal{O}(A)$ is a holomorphic function

---

[3] The gauge action $\mathrm{e}^{-S_g(A)}$ is typically an exponential and $\det D(A)$ is a polynomial. The action in this case is the product of two holomorphic functions and is therefore also holomorphic.

of A then

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} DA\ \mathrm{e}^{-S(A)} \mathcal{O}(A)}{\int_{\mathbb{R}^N} -DA\ \mathrm{e}^{-S(A)}} = \frac{\int_{\mathcal{M}} DA\ \mathrm{e}^{-S(A)} \mathcal{O}(A)}{\int_{\mathcal{M}} DA\ \mathrm{e}^{-S(A)}}\ . \tag{1.26}$$

Requiring that observables be holomorphic is also hardly a restriction. Bosonic observables are quite generally polynomials of the bosonic fields of the theory, and being polynomials they are holomorphic. Fermionic observables lead to holomorphic integrands as well. This is most readily demonstrated by an example. Consider the density $\langle n \rangle = \langle \bar{\psi}\gamma^0\psi \rangle$. The integration over fermions with the insertion of $\bar{\psi}\gamma^0\psi$ can be carried out by expanding the very large number of terms in the Grassmann integral:

$$\int D\bar{\psi} D\psi\ \bar{\psi}\gamma^0\psi\ \mathrm{e}^{-\sum_{xy} \bar{\psi}_x D_{xy} \psi_y} \tag{1.27}$$

Very few terms in the expansion will contribute to the Grassmann integral, and those that survive will be products of elements of the Dirac matrix. But the elements of the Dirac matrix are generally well behaved functions like $\mathrm{e}^{iA_\mu(x)}$ [4] and constant factors such as the mass. A finite product of such functions is holomorphic, so it's clear that the integrand of the path integral after integrating out the fermions is a holomorphic function. By the same logic, the integrand of the path integral after integrating out the fermions for any operator insertion of products of fermionic fields will behave in this way.

Now, given the freedom to deform the path integral allotted by Cauchy's theorem, we tame the sign problem by integrating over carefully chosen manifolds where the phase fluctuations $\mathrm{e}^{-i\mathrm{Im}\ S}$ which cause the sign problem are damped. To see how this might work, notice that while observables are unchanged by a deformation process, the average sign is changed. This is because the average sign is

$$\langle \mathrm{e}^{-i\mathrm{Im}\ S} \rangle_{\mathrm{Re}\ S} = \frac{\int DA\ \mathrm{e}^{-S}}{\int DA\ \mathrm{e}^{-\mathrm{Re}\ S}}\ . \tag{1.28}$$

Since the denominator is the integral of a non-holomorphic function, the average sign is manifold dependent. The idea is to make this quantity as large as possible by choosing "good" integration contours. There is more than one way to generate manifolds which increase the average sign. In this thesis we will concentrate on two methods.

The first method I will cover is called the *holomorphic gradient flow*. In this method, the original domain of integration is evolved by a first order differential equation to a manifold which has gentle phase fluctuations. This method evolved out of an initial proposal by Witten [24, 25] to perform path integrals over special contours in the complexified field space called "Lefschetz thimbles" with the intent to answer certain questions in Chern-Simons theories. We will not say much about Lefschetz thimbles in this thesis, but a few comments are in order to put the holomorphic gradient flow in context.

To begin, Lefschetz thimbles are the multi-dimensional analog of stationary phase contours [26, 27] and they are defined as follows. Given a holomorphic function $S : \mathbb{C}^N \to \mathbb{C}$ (call it the action), for every extrema z^* of the action (i.e. a point where $\partial S/\partial z_i(z^*) = 0$) there is an associated thimble. The thimble associated with z^* is defined as the union of all curves in complex field space which satisfy the following differential equation:

$$\frac{dz_i}{ds} = \left(\frac{\partial S}{\partial z_i}\right)^* , \tag{1.29}$$

and which satisfy $\lim_{s\to-\infty} z(s) = z^*$. Geometrically, the thimble associated to z^* is the union of all trajectories which emanate from z^* and satisfy Eq. (1.29). The imaginary part of the action is constant on a thimble. To see this, consider an arbitrary flow trajectory on a thimble. Then the change in the action along this flow trajectory is

$$\frac{d}{ds}S(z(s)) = \frac{\partial S(z(s))}{\partial z_i}\frac{dz_i}{ds} = \frac{\partial S(z(s))}{\partial z_i}\frac{\partial S(z(s))}{\partial z_i}^* > 0\ . \tag{1.30}$$

Since the change in action is strictly positive, the imaginary part does not change, and therefore we see that the imaginary part of the action is constant on a thimble. We see that a function of the form e^{-S} has a constant phase on a thimble, and it is in this sense that a thimble is the multi-dimensional generalization of paths of stationary phase.

What makes Lefschetz thimbles useful in computations is the so-called "thimble decomposition", which states the following: Given a holomorphic action, the following equality holds [28]

$$\int_{\mathbb{R}^N} dz\ e^{-S} = \sum_i n_i \mathrm{e}^{-i\mathrm{Im}\ S(z_i^*)} \int_{\mathcal{T}_i} dz\ \mathrm{e}^{-\mathrm{Re}\ S}\ , \tag{1.31}$$

where z_i^* are the extrema of the action, $\mathcal{T}_i$ is the thimble attached to z_i^* and n_i are integers called "intersection numbers"[4]. This means that the path integral can be decomposed into a sum of integrals over thimbles. Since the phase $\mathrm{e}^{-i\mathrm{Im}\ S}$ is constant a each thimble, ostensibly the sign problem is alleviated by an integration over thimbles.

The idea of using thimbles as a practical tool to alleviate the sign problem in lattice field theory originated in the work of Cristoforetti, Di Renzo and Scorzato [29]. This led to a rapid development of the "Lefschetz thimble method" for taming the sign problem in lattice field theory [30–47]. There are difficulties in using the Lefschetz thimble method in practice however. The main difficulty in applying the method is that many thimbles can contribute to the path integral. There doesn't

---

[4] The n_i give information about the weight a given thimble contributes to the path integral. Some thimbles have $n_i = 0$ and do not contribute to the path integral.

currently exist a generic way to find all extrema to which the thimbles are attached, nor is there a generic method for computing the intersection numbers n_i, so if many thimbles contribute to the path integral these issues become problematic. It is in striving to overcome these difficulties that the holomorphic gradient flow was formulated. The holomorphic gradient flow is a way to deform the original domain of integration to a manifold which approximates the set of thimbles which contribute to the theory without having to know the position of the extrema or the intersection numbers of the thimbles.

The second method for taming the sign problem I will describe in this thesis is called the *sign-optimized manifold method* [48]. Also based on Cauchy's theorem, the sign-optimized manifold method is a different way to deform path integrals to tame the sign problem. Rather than deforming the domain of integration by a flow evolution, one proposes a family of manifolds at the outset and then computes observables on an optimal member of the proposed family. To be quantitative, recall that the average sign is

$$\langle \sigma \rangle \equiv \langle \mathrm{e}^{-i\mathrm{Im}\ S} \rangle_{\mathrm{Re}\ S} = \frac{\int DA\ \mathrm{e}^{-S}}{\int DA\ \mathrm{e}^{-\mathrm{Re}\ S}} \tag{1.32}$$

and being an integral of a non-holomorphic function, changes from one manifold to the next. The goal of the sign-optimized manifold method is to find the manifold which maximizes the average sign. More procedurally, if we denote the family of manifolds as $\mathcal{M}(\lambda)$, then we want to find a set of λ for which $\nabla_\lambda \langle \sigma \rangle = 0$. The form of the proposed manifolds can vary, provided the stipulations of Cauchy's theorem are met, and knowledge of the particularities of the system of interest help to guide the proposal. In our formulation, the optimal manifold is found by stochastic gradient ascent in manifold space. This method has been used to solve the sign-problems in theories of fermions at finite density [22, 48].

This thesis is organized as follows. In Ch. 2 we will define the holomorphic gradient flow and describe how to integrate over a complex manifold with Monte Carlo methods. In Ch. 3 we will apply the methods developed to a toy fermionic model with a sign problem as a proof of principle. In Ch. 4 we will explain several important optimizations developed to improve the speed of sampling on complex manifolds. Utilizing these optimizations, we proceed to Ch. 5 where we compute the thermodynamic properties of a field theory of fermions with a sign problem. In Ch. 6 we explain the sign-optimized manifold method and apply it to another theory of fermions. Appendices with supplemental information can be found in Ch. 8.

We conclude this introduction with a few comments. The first is that the holomorphic gradient flow and the sign optimized manifold method are not the only approaches to taming the sign problem. The sign problem occurs in various fields of physics, and consequently many approaches have been developed to tame it. I attempt to compile a reasonably comprehensive list of approaches to the sign problem in lattice field theory here.

One popular approach is the Complex Langevin method, which is a generalization of the Langevin method to complex field variables, and has been applied to a number of models from cold atoms to QCD [49–59].The Complex Langevin method is straightforward and is numerically inexpensive. however, certain conditions need to be met by the theory of interest for the the method to converge to correct results [60,61]. The relatively new "path-optimization method" is another procedure for dealing with the sign problem. In this approach, a cost function which quantifies the sign problem is defined and minimized over a set of integration manifolds. This approach has been applied to a number of toy models [62–65].

Dual variables is another approach to the sign problem. Here the degrees of freedom of the partition function are changed such that, in the new variables the partition function is a sum of positive weights. When such a dual formulation can be found, the sign problem is manifestly absent. This approach has been applied a wide number of systems [66–76]. Another approach to the sign problem are the so-called "cluster algorithms", in particular the "meron-cluster" alogrithm, in which terms in the partition function are grouped into sub-sums over flips of "clusters", with the sub-sums contributing non-negatively to the partition function [77–84]. The fermion bag approach is yet another approach to dealing with the sign problem in fermionic systems [85–91]

Finally, in this document we will be solely concerned with computing the thermodynamic properties of relativistic theories of fermions at finite density. This narrow scope is due to my interest in such systems. However, a host of other systems also suffer from the sign problem. This includes bosons at finite density, non-relativistic systems as well as systems evolving in real-time. Cauchy's theorem and the holomorphic gradient flow can equally well be applied to these other systems, and it has been [19, 20, 92, 93]. The methods developed here are completely generic and can be applied to any theory with a complex action.

Chapter 2

The Holomorphic Gradient Flow

2.1 Flowing Points and Volumes

The holomorphic gradient flow is a way to generate a manifold in the complexified field space which satisfies the stipulations of Cauchy's Theorem [18]. To set notation, let us denote our fields which live on a hypercubic spacetime lattice with N sites as $(\phi_1, \phi_2, ..., \phi_N)$ and let us denote the action governing the system of interest as S. The fields ϕ_x can have internal degrees of freedom, however we will suppress these indices. We will concern ourselves with theories having complex actions, which are typical for systems at finite density. The holomorphic gradient flow a mapping between points on the original path integration domain and a submanifold of the complexified field space. The mapping is generated by the following first order differential equation

$$\frac{d}{dt}\phi(t) = \left(\frac{\partial}{\partial\phi}S(\phi(t))\right)^* . \tag{2.1}$$

Here the star means complex conjugate and t is a parameter with no physical relevance whose purpose is simply to parameterize the extent of the flow. We will leave the spacetime index of ϕ implicit for most of the discussion for neatness.

When clairity demands I will write indices. In components Eq. (2.1) reads

$$\frac{d}{dt}\phi_x(t) = \left(\frac{\partial}{\partial\phi_x}S(\phi(t))\right)^* \tag{2.2}$$

One sees from the definition of the holomorphic gradient flow that, if the action is complex, the fields evolve in a complex direction. Therefore, given a real field ϕ as an initial condition, the holomorphic gradient flow will evolve fields in the real plane $\mathbb{R}^N$ to fields into the complex plane $\mathbb{C}^N$.

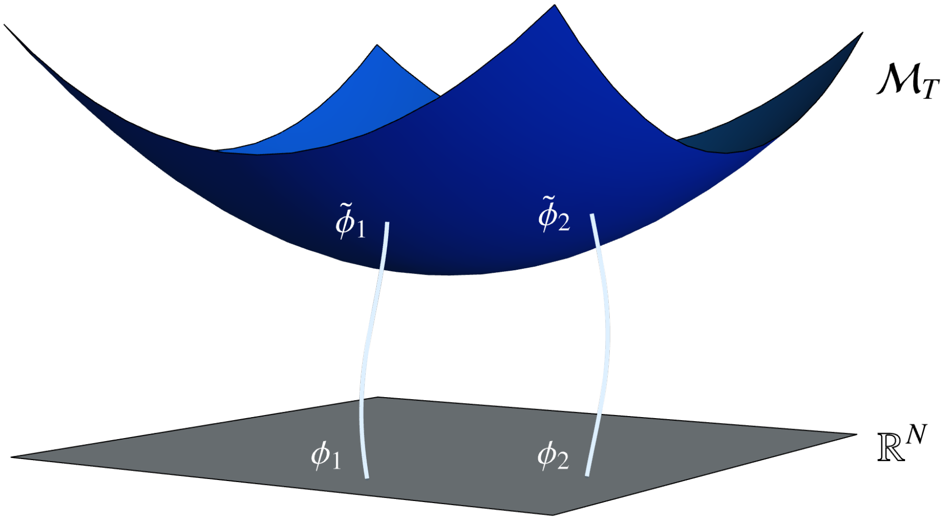

Figure 2.1: A schematic depiction of the deformation process. Arbitrary points $\phi \in \mathbb{R}^N$ evolve according to the holomorphic gradient flow, deforming the flat manifold $\mathbb{R}^N$ into the curved manifold $\mathcal{M}_T \subset \mathbb{C}^N$.

Now consider subjecting all points in the original domain of integration to a certain amount of flow time T. Let us suppose that $\phi \in \mathbb{R}^N$. Then each individual point ϕ will evolve to a new point $\tilde{\phi}$ in the complex plane which has the explicit expression

$$\tilde{\phi}(\phi) = \phi + \int_0^{T_{\text{flow}}} dt \, \left(\frac{\partial S}{\partial \phi}\right)^* \tag{2.3}$$

The set of all evolved points $\tilde{\phi}$ forms a new manifolds The facts of the past few paragraphs can be summarized by what we shall call the "flow map". For a selected flow-time T, the flow map $\Phi_T : \mathbb{R}^N \to \mathbb{C}^N$ is the function which takes a real valued field $\phi \in \mathbb{R}^N$ to the complex valued field $\tilde{\phi} \in \mathbb{C}^N$ obtained by applying the holomorphic gradient flow for an amount of time T:

$$\Phi_T(\phi) = \phi + \int_0^{T} dt \, \left(\frac{\partial S}{\partial \phi}\right)^* \tag{2.4}$$

Applying the flow map Φ_T to the original domain of integration $\mathbb{R}^N$ generates a complex manifold $\mathcal{M}_T \subset \mathbb{C}^N$. Symbolically, $\Phi(\mathbb{R}^N) = \mathcal{M}_T$.

Having generated a manifold by flowing the real plane, let us understand how to integrate on such a surface. Any observable computed in the grand-canonical ensemble has the path integral expression

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\phi) \mathrm{e}^{-S(\phi)}}{\int_{\mathbb{R}^N} D\phi \; \mathrm{e}^{-S(\phi)}} \tag{2.5}$$

Here I have denoted explicitly the domain of integration $\mathbb{R}^N$. Because a flowed manifold $\mathcal{M}_T$ can be generated from $\mathbb{R}^N$ without violating the stipulations of Cauchy's Theorem (I still have yet to prove this), one has by Cauchy's Theorem the following equality

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\phi) \mathrm{e}^{-S(\phi)}}{\int_{\mathbb{R}^N} D\phi \; \mathrm{e}^{-S(\phi)}} = \frac{\int_{\mathcal{M}_T} D\phi \; \mathcal{O}(\phi) \mathrm{e}^{-S(\phi)}}{\int_{\mathcal{M}_T} D\phi \; \mathrm{e}^{-S(\phi)}} \tag{2.6}$$

It is evident that any flowed manifold gives the same value for observables.

Typically a flowed manifold $\mathcal{M}_T$ is a complicated surface in the complex plane which cannot be written in a closed form. In spite of this difficulty, it is straightforward to describe how to integrate on such manifolds [18]. First, to integrate on a flowed manifold $\mathcal{M}_T$, it is necessary pick a parameterization of the surface. Since the flowed points in $\tilde{\phi} \in \mathcal{M}_T$ are in one-to-one correspondence with unflowed points $\phi \in \mathbb{R}^N$, it is clear that $\mathcal{M}_T$ can be parameterized by N real numbers, i.e. it is a manifold of N real dimensions. Let us call our coordinates ϕ. That is, for every $\tilde{\phi} \in \mathcal{M}_T$ there is an N dimensional array of real numbers ϕ such that $\tilde{\phi} = \tilde{\phi}(\phi)$. Keep in mind that even though both the coordinate system and the original domain of integration are parameterized by $\mathbb{R}^N$, they do not represent the same geometrical objects. The original domain of integration *is* $\mathbb{R}^N$ while the coordinates are surrogate quantities which represent points on $\mathcal{M}_T$. In coordinates, observables take the form

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathcal{M}_T} D\phi \ \mathcal{O}(\phi) \mathrm{e}^{-S(\phi)}}{\int_{\mathcal{M}_T} D\phi \ \mathrm{e}^{-S(\phi)}} = \frac{\int_{\mathbb{R}^N} D\phi \ \det J(\phi) \mathcal{O}(\tilde{\phi}(\phi)) \mathrm{e}^{-S(\tilde{\phi}(\phi))}}{\int_{\mathbb{R}^N} D\phi \ \det J(\phi) \mathrm{e}^{-S(\tilde{\phi}(\phi))}} \tag{2.7}$$

where $J(\phi)$ is the Jacobian matrix associated with the choice of coordinates used to describe the manifold and $\det J(\phi)$ is the volume element. It is obvious by the last equality that to compute the path integral one must be able to compute the Jacobian $\det J(\phi)$. We will find that computing the Jacobian is the most computationally expensive task when integrating on flowed manifolds.

We choose the following set of coordinates: the coordinate of $\tilde{\phi}$ is its pre-image under the flow [18]. In this language, the flow map Φ_T is the map which takes coordinates to points on the manifold and we can use Φ_T to compute the Jacobian associated with this choice of coordinates. The integration measure satisfies

$$d\tilde{\phi} = d\phi \frac{d\tilde{\phi}}{d\phi} \tag{2.8}$$

We therefore see that $J(\phi) = \frac{\partial \tilde{\phi}}{\partial \phi}$ and since $\tilde{\phi}(\phi) = \Phi_T(\phi)$ one has

$$J(\phi) = \frac{\partial}{\partial \phi} \Phi_T(\phi) \tag{2.9}$$

Writing $J(\phi)$ is not possible in closed form for an arbitrary flow map. Fortunately in the case of the holomorphic gradient flow $J(\phi)$ can be written, albeit implicitly, as the solution to a differential equation [94].

To derive this differential equation, note that Eq. (2.9) suggests that to look at nearby points ϕ and ϕ' on the coordinate manifold and examine the difference between their images under the flow. Define the difference between two nearby flowed points as $\delta\phi(t) = \phi'(t) - \phi(t)$. Although it may not be possible compute $\delta\phi(t)$ explicitly, it is possible to compute how $\delta\phi(t)$ changes in a short amount of flow time. Expanding $\delta\phi(t)$

$$\begin{aligned}
\delta\phi(t) \equiv \phi'(t) - \phi(t) &= \Big\{\phi'(0) + \Delta t \frac{d\phi'(t)}{dt}|_{t=0}\Big\} - \Big\{\phi(0) + \Delta t \frac{d\phi(t)}{dt}|_{t=0}\Big\} + \mathcal{O}(\Delta t^2) \\
&= \delta\phi(0) + \Delta t \Big\{\frac{d\phi'(t)}{dt}|_{t=0} - \frac{d\phi(t)}{dt}|_{t=0}\Big\} + \mathcal{O}(\Delta t^2) \\
&= \delta\phi(0) + \Delta t \Big\{\frac{\partial S(\phi)}{\partial \phi}|_{\phi=\phi'(0)} - \frac{\partial S(\phi)}{\partial \phi}|_{\phi=\phi(0)}\Big\}^* + \mathcal{O}(\Delta t^2) \\
&= \delta\phi(0) + \Delta t \Big\{\frac{\partial^2 S}{\partial \phi^2} \delta\phi(0)\Big\}^* + \mathcal{O}(\Delta t^2)
\end{aligned}$$

Taking the limit as $\Delta t \to 0$ one finds that, along a flow trajectory $\phi(t)$, a vector evolves as

$$\frac{d}{dt}\delta\phi(t) = \Big(H(\phi(t))\delta\phi\Big)^* \tag{2.10}$$

where we have called the Hessian, the matrix of second derivatives $H_{xy} = \frac{\partial^2 S}{\partial \phi_x \partial \phi_y}$.

It follows then that matrices evolve as

$$\frac{d}{dt}J(t) = \Big(H(\phi(t))J(t)\Big)^* \tag{2.11}$$

One can see this by recognizing that a matrix is a row vector of vectors:

$$J = \Big[\vec{v}_1, ..., \vec{v}_N\Big] \tag{2.12}$$

Therefore

$$\frac{d}{dt}J = \Big[\tfrac{d}{dt}\vec{v}_1, ..., \tfrac{d}{dt}\vec{v}_N\Big] \tag{2.13}$$

$$= \Big[H\vec{v}_1, ..., H\vec{v}_N\Big]^* \tag{2.14}$$

$$= \Big(H\Big[\vec{v}_1, ..., \vec{v}_N\Big]\Big)^* \tag{2.15}$$

$$= \Big(HJ\Big)^* \tag{2.16}$$

We can use this finding to compute $\det J(\phi)$. Note that $\det J(\phi)$ the ratio of the volume spanned by a set of vectors after the flow map to the volume spanned by the set of vectors before the flow map. If we take the original set of vectors to be the canonical basis on $\mathbb{R}^N$, then $\mathbb{1}$ is the matrix before the flow map. Applying the flow map, and using Eq. (2.11), the Jacobian matrix is given by

$$J(\phi) = \mathbb{1} + \int_0^T dt\ \Big(H(\phi(t))J(\phi(t))\Big)^* \tag{2.17}$$

where $\phi(t)$ is evaluated on the flow trajectory with initial condition ϕ.

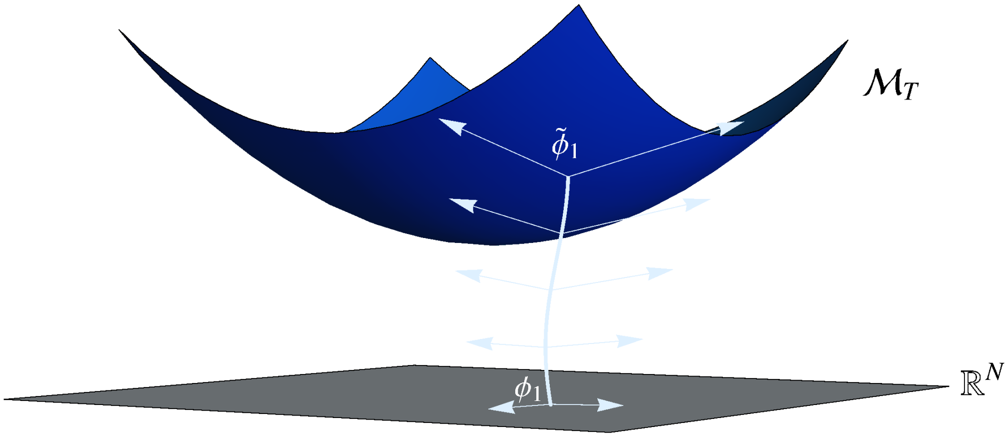

Figure 2.2: Above is a schematic picture of the evolution of a pair of displacement vector as a function of flow time.

Computing $J(\phi)$ for a general theory is hard because in order to do so one must solve a coupled point-flow/matrix-flow evolution which typically must be done numerically. This is always an expensive procedure, but the expense differs greatly in the presence and absence of fermions. This is not surprising; the inclusion of fermions always increases the cost of lattice calculations due to the presence of the fermion determinant. It is instructive, however, to estimate the cost of the holomorphic gradient flow with and without fermions.

In the absence of fermions, lattice actions are typically local and of the form

$$S(\phi) = \sum_x f_x(\phi) \tag{2.18}$$

where f_x is a function of ϕ_x and its nearest neighbors. Therefore computing $\frac{\partial S}{\partial \phi_x} \sim \mathcal{O}(1)$. This means that a flow trajectory costs $\mathcal{O}(N \times N_{step})$ where N is the number of lattice sites and N_{step} is the number of discrete time steps in the holomorphic gradient flow evolution. Similarly, the cost of computing the Jacobian for a bosonic theory is $\mathcal{O}(N^2 \times N_{step})$. In the presence of fermions, however, actions get a contribution from the fermion determinant

$$S(\phi) = \sum_x f_x(\phi) - \ln \det D(\phi) \tag{2.19}$$

and the cost of the holomorphic gradient flow is increased. To find the cost of flowing with fermions, note that for a single time step and for a single lattice site

$$\frac{d\phi'_x(t)}{dt} \sim f'_x(\phi) - \mathrm{tr}(D^{-1}(\phi)\frac{\partial D(\phi)}{\partial \phi_x}) \tag{2.20}$$

The first term has a cost of $\mathcal{O}(1)$ but the second term is more expensive. Typically $D(\phi)$ are sparse matrices [3], and $\frac{\partial D}{\partial \phi_x}$ has $\mathcal{O}(1)$ non-zero entries, independent of the number of lattice sites. In particular $\partial D/\partial \phi_x$ is "ultra-local", meaning that

$$\left(\frac{\partial D}{\partial \phi_x}\right)_{x'y'} \sim \delta_{xx'}\delta_{xy'} \tag{2.21}$$

This imples that for a single time step, to flow all points costs a matrix inversion. This means to computing a single flow trajectory costs $\mathcal{O}(N^3 \times N_{step})$, which is N^2 more expensive than the bosonic case. Evolving the Jacobian is still more expensive. To compute a single time step of the discretized evolution of

$$\frac{dJ}{dt} = \left(HJ\right)^* \tag{2.22}$$

it is necessary to compute the matrix product HJ. To understand the cost of this multiplication, first note that $H_{xy} \sim D^{-1}_{xy} D^{-1}_{yx}$. Thus to compute the hessian, one must first pay $\mathcal{O}(N^3)$ to invert D and then $\mathcal{O}(N^2)$ more to assign each element of the hessian. Finally, one must multiply matrices to compute HJ, and the cost of multiplying matrices is $\mathcal{O}(N^3)$. All together, a single time step has a cost of $\mathcal{O}(N^8)$. Consequently flowing the Jacobian costs $\mathcal{O}(N^8 \times N_{step})$. It is sometimes possible to construct more sophisticated algorithms to flow the Jacobian which cost less than the straightforward algorithm detailed here [23].

2.2 Monte Carlo on Flowed Manifolds

Using the holomorphic gradient flow to deform the domain of integration of the path integral, we are left with the task of evaluating

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} D\phi \ \det J(\phi) \mathcal{O}(\tilde{\phi}(\phi)) \mathrm{e}^{-S(\tilde{\phi}(\phi))}}{\int_{\mathbb{R}^N} D\phi \ \det J(\phi) \mathrm{e}^{-S(\tilde{\phi}(\phi))}} \ . \tag{2.23}$$

To compute $\langle \mathcal{O} \rangle$ we will use Monte Carlo methods. I will assume the reader is accustomed with the basic ideas of Markov Chain Monte Carlo and those less familiar with may find [3,4,12] to be useful references.

To begin a Monte Carlo evaluation of Eq. (2.23) we first define the effective action

$$S_{eff}(\phi) = S(\tilde{\phi}(\phi)) - \ln \det J(\phi) \ , \tag{2.24}$$

which includes the Jacobian. Note that $S(\phi)$ includes a fermion determinant if the theory has fermions. Using the effective action we can rewrite observables in the reweighted form

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} D\phi \ \mathcal{O}(\tilde{\phi}(\phi)) \mathrm{e}^{-S_{eff}(\tilde{\phi}(\phi))}}{\int_{\mathbb{R}^N} D\phi \ \mathrm{e}^{-S_{eff}(\tilde{\phi}(\phi))}} \tag{2.25}$$

Separating the real and imaginary parts of the action, we can rewrite

$$\begin{aligned}
\langle \mathcal{O} \rangle &= \frac{\int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\tilde{\phi}(\phi))\mathrm{e}^{-S_{eff}(\tilde{\phi}(\phi))}}{\int_{\mathbb{R}^N} D\phi \; \mathrm{e}^{-S_{eff}(\tilde{\phi}(\phi))}} \\
&= \frac{\int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\tilde{\phi}(\phi))\mathrm{e}^{-i\mathrm{Im}S_{eff}(\tilde{\phi}(\phi))}\mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))}}{\int_{\mathbb{R}^N} D\phi \; \mathrm{e}^{-i\mathrm{Im}S_{eff}(\tilde{\phi}(\phi))}\mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))}} \\
&= \frac{\int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\tilde{\phi}(\phi))\mathrm{e}^{-i\mathrm{Im}S_{eff}(\tilde{\phi}(\phi))}\mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))} / \int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\tilde{\phi}(\phi))\mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))}}{\int_{\mathbb{R}^N} D\phi \; \mathrm{e}^{-i\mathrm{Im}S_{eff}(\tilde{\phi}(\phi))}\mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))} // \int_{\mathbb{R}^N} D\phi \; \mathcal{O}(\tilde{\phi}(\phi))\mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))}} \\
&= \frac{\langle \mathcal{O}\mathrm{e}^{-i\mathrm{Im}S_{eff}} \rangle_{\mathrm{Re}\ S_{eff}}}{\langle \mathrm{e}^{-i\mathrm{Im}S_{eff}} \rangle_{\mathrm{Re}\ S_{eff}}}
\end{aligned}$$

where $\langle \ \cdot \ \rangle_{\mathrm{Re}\ S_{eff}}$ is an average with respect to the probability distribution

$$p(\phi) = \frac{e^{-\mathrm{Re}\ S_{eff}(\phi)}}{\int_{\mathbb{R}^N} D\phi \; \mathrm{e}^{-\mathrm{Re}S_{eff}(\tilde{\phi}(\phi))}} \tag{2.26}$$

As stated in the Introduction, this re-writing of observables as a ratio of observables with fields distributed as Eq. (2.26) is called “reweighting”. This is because absorbing on the fluctuating phase $\mathrm{e}^{-i\mathrm{Im}S_{eff}}$ alters the weight of a given configuration.

Our Monte Carlo strategy for integrating on flowed manifolds is to sample fields which live on the coordinate manifold according to $p(\phi)$, and then compute observables with the reweighting technique. There are many algorithms to do this sampling and we elect to use the Metropolis algorithm [95]. To begin a flow time T is chosen; this defines the manifold that will be integrated over. Then the following steps are iterated:

1. Begin with a configuration $\phi^{(i)}$ on the coordinate manifold.

2. Evolve $\phi^{(i)}$ according to the holomorphic gradient flow for time T and call the evolved point $\tilde{\phi}^{(i)}$. With the flow trajectory in hand, evolve the Jacobian according to its flow equation Eq. (2.17).

3. Compute the effective action $S_{eff}(\phi^{(i)}) = S(\tilde{\phi}^{(i)}) - \ln \det J(\phi^{(i)})$ and save.

4. Propose a new configuation on the coordinate manifold $\phi^{(i+1)} = \phi^{(i)} + \delta\phi^{(i)}$ where $\delta\phi^{(i)}$ is a small variation.

5. Repeat steps $2-4$ with $\phi^{(i+1)}$

6. If $\mathrm{Re}S_{eff}(\phi^{(i+1)}) < \mathrm{Re}S_{eff}(\phi^{(i)})$ accept the proposed configuration immediately, otherwise accept the proposed configuration with probability

$$\mathrm{e}^{-(\mathrm{Re}S_{eff}(\phi^{(i+1)})-\mathrm{Re}S_{eff}(\phi^{(i)}))} \tag{2.27}$$

7. Repeat steps 1-6.

Fields sampled in this way will be distributed according to $p(\phi)$ in the limit that the number of configurations collected tends to infinity [18].

Chapter 3

Application 1: 0+1 Thirring Model

3.1 Continuum Theory and Lattice Formulation

We now apply the methods described in the previous chapter to a simple model of fermions. The idea here is to demonstrate that the program outlined in the last chapter indeed works. We will demonstrate the method by computing the properties of the finite density Thirring model in zero spatial dimensions. A few particularities arise in the analysis due to the fermion determinant, but the procedure described in the previous chapters is unchanged.

In the general case, the Thirring model at finite density in d dimensions is defined by the continuum Euclidean action

$$S = \int dt d^d x \Big\{ \overline{\psi}(\partial\!\!\!/ + m + \mu\gamma^0)\psi + \frac{g^2}{2}\big(\overline{\psi}\gamma^\mu\psi\big)^2 \Big\} \ , \tag{3.1}$$

where ψ is a Dirac fermion and γ^μ are Euclidean Dirac matrices satisfying $\{\gamma^\mu, \gamma^\nu\} = 2\mathbb{1}_{d\times d}\delta_{\mu,\nu}$. The Thirring model (at $m = 0$ and $\mu = 0$) was introduced quite some time ago, originally in $(1+1)$ dimensions [96], and is one of the few examples of an "exactly solvable" quantum field theory. Various solutions of the massless case are given here [97, 98].

In the $m \neq 0$ case the $S-$matrix has been explicitly computed [99] and it was shown by Coleman [100] that the Thirring model at zero density is equivalent to the sine-Gordon model, which, interestingly, is a theory of bosons rather than fermions. We choose to test our methods on the Thirring model both because it is a simple model which is well-studied, and because it is a theory of fermions, which ostensibly forces us to face the high cost of such calculations. Aside from its simplicity, however, there are other reasons for looking at this model. For example, there is reason to expect the finite density properties of the Thirring model to have some resemblance to the finite density properties of QCD [101], but from our perspective what's important is not the Thirring model's proximity to QCD, but rather the fact that it has a sign problem which can be made arbitrarily severe.

In this chapter we apply the holomorphic gradient flow to (0+1) version of the theory. Here the action reduces to

$$S = \int dt \Big\{ \overline{\psi}\big(\frac{d}{dt}\gamma^0 + m + \mu\gamma^0\big)\psi + \frac{g^2}{2}\big(\overline{\psi}\gamma^\mu\psi\big)^2 \Big\} \ . \tag{3.2}$$

To analyze this theory on the lattice, it pays to introduce an auxiliary vector field A_μ whose role is to decouple the four fermi interaction. By noting that

$$\int DA_\mu \mathrm{e}^{-(\frac{1}{2g^2}A_\mu^2 + iA_\mu(\overline{\psi}\gamma^\mu\psi))} = \text{const} \times \mathrm{e}^{-\frac{g^2}{2}(\overline{\psi}\gamma^\mu\psi)(\overline{\psi}\gamma^\mu\psi)} \ , \tag{3.3}$$

one sees that the action in terms of auxiliary fields reads (in d dimensions)

$$S = \int dt d^d x \Big\{ \frac{1}{2g^2}A_\mu^2 + \overline{\psi}(\not{\partial} + i\not{A} + m + \mu\gamma^0)\psi \Big\} \ . \tag{3.4}$$

In $(0+1)$ dimensions A_μ has only one component, A_0, which we will call ϕ to align with the notation of the previous chapter. We now put this theory on the lattice. In our calculations we simulate an action whose naive continuum limit is

Eq. (3.4); the lattice action reads

$$S = \frac{1}{\hat{g}^2}\sum_t \Big(1-\cos(\hat{\phi}_t)\Big) + \hat{m}\sum_t \overline{\psi}_t\psi_t + \frac{1}{2}\sum_t \Big(\overline{\psi}_t \mathrm{e}^{i\hat{\phi}_t+\hat{\mu}}\psi_{t+1} - \overline{\psi}_{t+1}\mathrm{e}^{-i\hat{\phi}_t-\hat{\mu}}\psi_t\Big) \tag{3.5}$$

where t indexes the time slices and ranges from 0 to $N-1$ [18]. In the action Eq. (3.5), ϕ_t are real numbers while $\psi_t, \bar{\psi}_t$ are Grassmann variables satisfying $\{\psi_t, \psi_{t'}\} = \{\psi_t, \bar{\psi}_{t'}\} = \{\bar{\psi}_t, \bar{\psi}_{t'}\} = 0$. Here, because we are interested thermodynamics, the Grassmann fields are anti-periodic in time, i.e. $\psi_{N+1} = -\psi_1$ and the auxiliary fields are periodic in time $\phi_{N+1} = \phi_1$. These boundary conditions are fixed by the fact that we are doing thermodynamics, and therefore taking a trace with fermionic/bosonic coherent states to compute $\mathrm{tr}\,\mathrm{e}^{-\beta(H-\mu N)}$ [3].

3.2 Deforming Compact Fields and Proving Sufficient Conditions

Notice that the action is a function of a compact variable in our formulation, i.e. $S(\hat{\phi}_t) = S(\hat{\phi}_t + 2\pi)$ for any t. This means that our original domain of integration is not $\mathbb{R}^N$, but $(S_1)^N$, that is N copies of the circle S_1. This change in the domain of integration does not change the analysis in any fundamental way. Let us demonstrate that this is the case.

First, note that any element of $(S_1)^N$ is of the form $(\mathrm{e}^{i\phi_1}, ..., \mathrm{e}^{i\phi_N})$ where $\phi_i \in \mathbb{R}$. In our formulation, we choose to deform the argument ϕ_i into the complex plane, and therefore points in the complexified domain of integration are of the form

$$(\mathrm{e}^{i(\mathrm{Re}\phi_1 + i\mathrm{Im}\phi_1)}, ..., \mathrm{e}^{i(\mathrm{Re}\phi_N + i\mathrm{Im}\phi_N)}) \in (S_1 \times \mathbb{R})^N \ , \tag{3.6}$$

therefore the complexified field space is $(S_1 \times \mathbb{R})^N$. Our goal is to use the holomorphic gradient flow to deform the original domain of integration $(S_1)^N$ into a submanifold $\mathcal{M}_T$ of the complexified field space $(S_1 \times \mathbb{R})^N$. In order for this

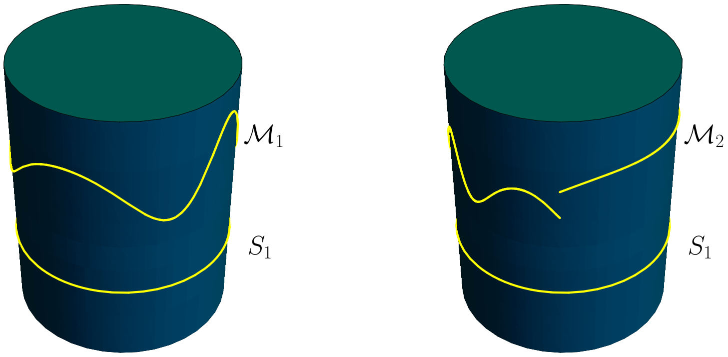

Figure 3.1: Here we depict possible manifold deformations for the simple case of a single variable. The blue cylinder represents the full complexified space $S_1 \times \mathbb{R}$. The original domain of integration is the yellow circle on the bottom, while possible deformed manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$ are shown above. In order for Cauchy's theorem to hold, it must be possible to deform S_1 to the final manifold smoothly. Therefore $\mathcal{M}_1$ can be used for integration while $\mathcal{M}_2$ cannot be used due to the discontinuity.

deformation to leave the path integral invariant, it must be the case that the stipulations of Cauchy's theorem hold for the deformation in question. The behavior of the integrand at the boundary must not be changed, and one must not cross any singularities of the integrand in the deformation process.

Since $(S_1)^N$ has no boundary (it's a compact manifold) the first assumption of Cauchy's theorem is satisfied trivially, as long as the new intgration contour can be be produced by a continuous deformation of the original domain of integration. It is important that no discontinuities are introduced in the deformation process; such deformations are not allowed. For two examples of deformations see Fig. (7.1).

Manifolds generated by the holomorphic gradient flow have no discontinuities. To see this, note that $\tilde{\phi}(\phi) = \tilde{\phi}(\phi + 2\pi\hat{e}_i)$ for any i, where $2\pi\hat{e}_i$ is a 2π shift in the i^{th} direction. This is because $\partial S/\partial\phi(\phi) = \partial S/\partial\phi(\phi + 2\pi\hat{e}_i)$, which is true because S is periodic. There cannot be any discontinuities in the manifold then, because if there were, it would be possible wrap 2π around this point and obtain two different $\tilde{\phi}$, contradicting the fact that $\tilde{\phi}(\phi) = \tilde{\phi}(\phi + 2\pi\hat{e}_i)$ for every i.

The second assumption of Cauchy's theorem is that no singularities are crossed during the deformation. No singularities are crossed during the flow in this model because there are no singularities in $(S_1 \times \mathbb{R})^N$ to cross. This is typically the case for field theories with fermions. To see this, note that after integrating out the fermions, the partition function takes the form

$$Z = \int D\phi \ \mathrm{e}^{-S_g(\phi)} \mathrm{det} D(\phi) \tag{3.7}$$

where $\mathrm{det}D(\phi)$ is the fermion determinant and $S_g(\phi) \equiv \frac{1}{\hat{g}^2} \sum_t \Big(1 - \cos(\hat{\phi}_t)\Big)$. It is clear that the integrand has no poles: the fermion determinant is a polynomial in the phases $\mathrm{e}^{i\hat{\phi}_t}$ and S_g is analytic. Furthermore, any fermion correlation function, such as the chiral condensate for example,

$$\langle \bar{\psi}_x \psi_x \rangle = Z^{-1} \int D\phi D\bar{\psi} D\psi \ \mathrm{e}^{-S_g(\phi)} \mathrm{e}^{-S_f(\bar{\psi},\psi)} \bar{\psi}_x \psi_x \tag{3.8}$$

also has an integrand without poles. This can be demonstrated by expanding the Grassmann integral. Let us consider a simple example; the generalization is straightforward. Consider a fermion action with two degrees of freedom

$$S_f(\bar{\psi}, \psi) = M_{11}\bar{\psi}_1\psi_1 + M_{12}\bar{\psi}_1\psi_2 + M_{21}\bar{\psi}_2\psi_1 + M_{22}\bar{\psi}_2\psi_2 \ . \tag{3.9}$$

where $M = M(\phi)$. Then the chiral condensate at variable 1 is given by

$$\begin{aligned} \langle \bar{\psi}_1 \psi_1 \rangle &= Z^{-1} \int D\phi \ \mathrm{e}^{-S_g(\phi)} \Big(\int D\bar{\psi} D\psi \mathrm{e}^{-S_f} \bar{\psi}_1 \psi_1 \Big) \\ &= Z^{-1} \int D\phi \ \mathrm{e}^{-S_g(\phi)} M_{22}(\phi). \end{aligned} \tag{3.10}$$

The entries of M are analytic functions, and therefore $\langle\bar{\psi}_1\psi_1\rangle$ is the integral of an analytic function. Therefore there are no singularities to cross when the path integral is deformed, so Cauchy's theorem holds and the integral doesn't change.

This might be surprising to some given that the chiral condensate also has another representation

$$\langle\bar{\psi}_1\psi_1\rangle = Z^{-1}\int D\phi\ \mathrm{e}^{-S_g(\phi)}\det D(\phi)\ D_{11}^{-1}(\phi) \tag{3.11}$$

and at points where $\det D = 0$, D^{-1} blows up. How can it be that the integrand has no poles? The answer is that the expression above is not always valid: the chiral condensate can only be written as Eq. (3.11) when the determinant is non-zero. On the other hand, Eq. (3.10) is always true and fermion correlators always have analytic path integrands. These two statements are not in contradiction because D^{-1} is always multiplied by $\det D$. As D^{-1} blows up, $\det D$ goes to zero in such a way that the limit of the path integrand exists as a zero of the determinant is approached. For clarity, suppose D is diagonal, then

$$M_{22}(\phi) = \det D(\phi) D_{11}^{-1}(\phi) = \frac{M_{11}(\phi)M_{22}(\phi)}{M_{11}(\phi)} \tag{3.12}$$

If we look at a region around a zero of the determinant, then nearby in field space both the left hand and right hand sides exist At the zero however, the right hand side does not exist, but the Grassmann integral still exists and is equal to $M_{22}(\phi)$. If one elects to calculate the path integrals by simulating ϕ with the fermions integrated out, which is typical, then the precise integrand of the path integral, say in the case of the chiral condensate, is

$$f(\phi) = \begin{cases} \mathrm{e}^{-S_g(\phi)} \det D(\phi) D_{11}^{-1}(\phi), & \text{when } \det D(\phi) \neq 0 \\ \lim_{\phi' \to \phi} \mathrm{e}^{-S_g(\phi')} \det D(\phi') D_{11}^{-1}(\phi'), & \text{when } \det D(\phi) = 0 \end{cases}$$

and so the integrand of the path integral is perfectly well-defined and finite. This is very much like what happens with the function $f(x) = \sin x/x$, which converges to 1 as $x \to 0$ even though f does not exist at zero.

Having demonstrated that we can apply Cauchy's theorem, we elect to deform the manifold in two steps [18]:

1. Shift all fields into the imaginary direction: $\hat{\phi}_i \to \hat{\phi}_i + i\alpha$. This shift causes the manifold to touch a saddle point of the action.
2. Apply the holomorphic gradient flow to the shifted manifold.

3.3 The Tangent Plane

The first step in our taming the sign problem of the Thirring model is a shift of the domain of integration in the purely imaginary direction. Let us understand why this might be advantageous for the purpose of taming phase fluctuations. Note that there is quite a bit of analysis that goes into showing this, however it is worth dwelling on because shifting the integration domain to the tangent plane is an inexpensive way to tame phase fluctuations.

To begin, consider an extrema ϕ^* of the action

$$\frac{\partial}{\partial \phi} S(\phi) \bigg|_{\phi = \phi^*} = 0 \tag{3.13}$$

Due to the fact that S is holomorphic, any such extrema is a saddle point, rather than a minimum or a maximum [102]. To see this, consider the case of a single complex variable: $S = S(z) = S_R(x,y) + iS_I(x,y)$ where $z = x + iy$. In order for S to be a holomorphic function it must satisfy the Cauchy-Riemann conditions. That is,

$$\frac{\partial S_R}{\partial x} = \frac{\partial S_I}{\partial y}, \quad \frac{\partial S_R}{\partial y} = -\frac{\partial S_I}{\partial x} \tag{3.14}$$

Taking derivatives of the Cauchy-Riemann conditions shows directly that $\nabla^2 S_R = 0$ and $\nabla^2 S_I = 0$. Now, changes of the action arise at second order when expanding about an extrema. Examining the Hessian of the real part:

$$H_R = \begin{bmatrix} \frac{\partial^2 S_R}{\partial x^2} & \frac{\partial^2 S_R}{\partial x \partial y} \\ \frac{\partial^2 S_R}{\partial x \partial y} & \frac{\partial^2 S_R}{\partial y^2} \end{bmatrix} \tag{3.15}$$

one finds that $\mathrm{tr}(H) = \nabla^2 S_R = 0$, and so it must be the case that the eigenvalues of H are either both zero, or both non-zero with opposite signs. Therefore any non-degenerate critical point has a direction where S_R increases and an orthogonal direction where S_R decreases, i.e it is a saddle point. The same analysis applies to S_I and therefore any non-degenerate critical point of a holomorphic function of a single variable is a saddle point.

To extend this analysis to a multivariate holomorphic function $S = S(z_1, ..., z_n)$ is straightforward. Recall that a holomorphic function is a function that satisfies the Cauchy-Riemann conditions in each variable (keeping all other variables fixed):

$$\frac{\partial S_R}{\partial x_i} = \frac{\partial S_I}{\partial y_i}, \quad \frac{\partial S_R}{\partial y_i} = -\frac{\partial S_I}{\partial x_i} \tag{3.16}$$

This means that $\nabla_i^2 S_R \equiv (\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2}) S_R = 0$ for each i (and similarly for S_I). Therefore, if we have an extrema $z^* = (z_1^*, ..., z_n^*)$ and a deviation δz_i in the i^{th} direction, we see that the Hessian at the critical point has a pair of eigenvalues with opposite signs. Since this applies to each variable, half of the eigenvalues of the Hessian at the critical point are positive and half are negative, i.e. the extrema is a saddle point in n dimensions.

We will now demonstrate by expanding the holomorphic gradient flow about an extrema of the action that there is a plane passing though the extremal point such that, at least close to the extrema, the phase fluctuations vanish. If the original domain of integration can be shifted and rotated into this plane then it can be used for integration and doing so tames the sign problem to a certain degree. After demonstrating the existence of this "tangent plane" generically, we will calculate explicitly the tangent plane for the Thirring model.

Let us call the extremum ϕ^* and expand

$$S(\phi) = S(\phi^*) + \frac{1}{2}\delta\phi_i H_{ij} \delta\phi_j + \mathcal{O}(\delta\phi^3) \tag{3.17}$$

where $H_{ij} = \frac{\partial^2}{\partial\phi_i \partial\phi_j} S(\phi)\Big|_{\phi=\phi^*}$. To quadratic order in the action, the holomorphic gradient flow is equal to

$$\frac{d\phi}{dt} = \Big(H\phi\Big)^* \tag{3.18}$$

This differential equation can be solved exactly. To do so, let $H = H_R + iH_I$ and $\phi = \phi_R + i\phi_I$. Then by separating the left and right hand sides into their real and imaginary parts, it is clear that

$$\frac{d}{dt}\begin{bmatrix}\phi_R \\ \phi_I\end{bmatrix} = \begin{bmatrix} H_R & -H_I \\ -H_I & -H_R\end{bmatrix}\begin{bmatrix}\phi_R \\ \phi_I\end{bmatrix} = \mathcal{H}\begin{bmatrix}\phi_R \\ \phi_I\end{bmatrix} \tag{3.19}$$

where we have defined the "super-hessian" $\mathcal{H} = \begin{bmatrix} H_R & -H_I \\ -H_I & -H_R\end{bmatrix}$. Note that $\mathcal{H}$ is a real symmetric matrix and is therefore diagonalizable. It is clear from the analysis in the previous paragraphs that $\mathcal{H}$ has paired positive/negative eigenvalues. This is because extrema of analytic functions are saddle points and a saddle point always has an increasing and a decreasing direction. Another way to see this is as follows: suppose that $\mathcal{H}\phi = \lambda\phi$ and define

$$\epsilon = \begin{bmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix} \tag{3.20}$$

It is simple to deduce that $\epsilon^2 = -\mathbb{1}$ and that $\epsilon\mathcal{H}\epsilon = \mathcal{H}$. Therefore

$$\mathcal{H}(\epsilon\phi) = -\epsilon\mathcal{H}\phi = -\lambda(\epsilon\phi) \tag{3.21}$$

So each eigenvector ϕ has another eigenvector $\epsilon\phi$ with an opposite eigenvalue. Taking the n eigenvectors with positive eigenvalue one can form a basis of the tangent space of the thimble attached to the extremum ϕ^*. To do this, denote every eigenvector of the super hessian by $\rho = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}$. Then the set $\{\rho_1 + i\rho_2\}$ is a basis of the tangent space of the thimble. The basis of the tangent space attached to an extremum which defines a thimble is sometimes called a set of "Takagi vectors".

Now for the point: moving the domain of integration to be the tangent space of the thimble aligns the integration domain with a thimble, at least in a region near the extrema. Therefore, since thimbles are surfaces of stationary phase, the tangent space at the extrema has small phase fluctuations and such a deformation is therefore useful for taming the sign problem. This statement holds for any path integral. Let us now carry out this procedure explicitly for the (0+1) Thirring Model.

The action

$$S = \frac{1}{\hat{g}^2}\sum_t \Big(1 - \cos(\hat{\phi}_t)\Big) + \hat{m}\sum_t \overline{\psi}_t\psi_t + \frac{1}{2}\sum_t \Big(\overline{\psi}_t \mathrm{e}^{i\hat{\phi}_t+\hat{\mu}}\psi_{t+1} - \overline{\psi}_{t+1}\mathrm{e}^{-i\hat{\phi}_t-\hat{\mu}}\psi_t\Big) \tag{3.22}$$

presumably has many extrema, most of which are very difficult to find. One way of attacking such a complicated problem is to search for extrema with a particular form, rather than to search for a generic one. To that end let us look in the sub-space of field configurations that are constant in spacetime, that is $\hat{\phi}_t = \phi$ for every t. For constant fields the action Eq. (3.22) can be diagonalized in momentum space. Using $\psi_t = \sum_\omega e^{-i\omega t}\psi_\omega$ where $\omega = \frac{(2k+1)\pi}{N_t}$, $k = 0, ..., N_t - 1$ then we find upon integrating out the fermions

$$S = \frac{1}{\hat{g}^2} N_t \Big(1 - \cos(\phi)\Big) - \sum_\omega \ln\Big(\mathrm{e}^{i\phi+\hat{\mu}-i\omega} - \mathrm{e}^{-i\phi-\hat{\mu}+i\omega}\Big) \tag{3.23}$$

Consequently the derivative is given by

$$\frac{\partial S}{\partial \phi} = \sum_\omega \Big[\frac{\sin\phi}{\hat{g}^2} - \cot(\phi - \omega - i\mu)\Big] \tag{3.24}$$

There is a point with purely imaginary ϕ where the derivative vanishes. The position of this point is a function of the parameters of the model $\hat{g}, \hat{\mu}$. An example calculation is shown in Fig. (3.2) where the extrema of the action is located at a constant field configuration and $\phi^* = i0.368...$

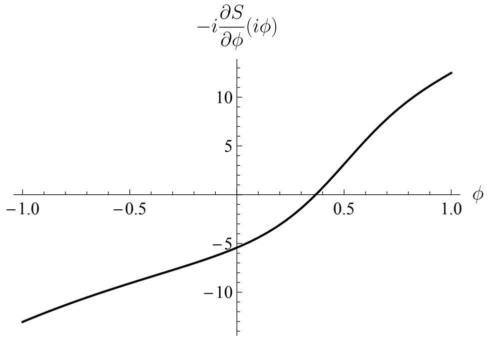

Figure 3.2: Here we show the derivative of the action evaluated along the imaginary axis for $\hat{g} = 1.0$ and $\hat{\mu} = 0.5$ on a lattice with $N_t = 6$.

It is a more challenging task to compute the hessian. Taking derivatives one finds

$$H_{tt'} = \frac{\partial^2 S}{\partial\phi_t \partial\phi_{t'}} = \delta_{tt'}\frac{\cos\phi_t}{\hat{g}^2} + \mathrm{tr}\left(D^{-1}\frac{\partial D}{\partial\phi_t}D^{-1}\frac{\partial D}{\partial\phi_{t'}}\right) - \mathrm{tr}\left(D^{-1}\frac{\partial^2 D}{\partial\phi_t\partial\phi_{t'}}\right) \tag{3.25}$$

To calculate these derivatives it is necessary to have the fermion matrix handy. Here it is:

$$D_{tt'} = \delta_{tt'}\hat{m} + \frac{1}{2}\Big(\big\{\delta_{t',t+1}\mathrm{e}^{i\phi_t+\hat{\mu}} - \delta_{t',1}\delta_{t,N}\mathrm{e}^{i\phi_N+\hat{\mu}}\big\} - \big\{\delta_{t',t-1}\mathrm{e}^{-i\phi_{t-1}-\hat{\mu}} - \delta_{t',N}\delta_{t,1}\mathrm{e}^{-i\phi_N-\hat{\mu}}\big\}\Big) \tag{3.26}$$

Note that the anti-periodic boundary conditions are the source of the $\delta_{t',1}\delta_{t,N}$ type terms. These boundary terms are a bit unruly, so for the remainder of the discussion I will render these terms implicit by agreeing to always adhere to the following:

- I will take the delta functions to be periodic in both arguments, i.e. $\delta_{t,t'} = \delta_{t+N,t'} = \delta_{t,t+N'}$
- Any input field $\phi = (\phi_1, ..., \phi_N)$ is to be mapped to $\phi = (\phi_1, ..., \phi_N + \pi)$ before being plugged into the dirac matrix. Since fields appear in the action as $\mathrm{e}^{i\phi}$ this mapping takes $\mathrm{e}^{i\phi_N} \longrightarrow -\mathrm{e}^{i\phi_N}$, enforcing anti-periodicity.

With these stipulations, the implicit form of the dirac matrix is

$$D_{tt'} = \delta_{tt'}\hat{m} + \frac{1}{2}\left(\delta_{t',t+1}\mathrm{e}^{i\phi_t+\hat{\mu}} - \delta_{t',t-1}\mathrm{e}^{-i\phi_{t-1}-\hat{\mu}}\right) \tag{3.27}$$

Taking the second derivative we find:

$$\left(\frac{\partial^2 D}{\partial\phi_a\partial\phi_b}\right)_{tt'} = -\frac{1}{2}\left(\delta_{t,a}\delta_{t,b}\delta_{t',t+1}\mathrm{e}^{i\phi_t+\mu} - \delta_{t-1,a}\delta_{t-1,b}\delta_{t',t-1}\mathrm{e}^{-i\phi_{t-1}-\mu}\right)\ . \tag{3.28}$$

Thus,

$$\mathrm{tr}\left(D^{-1}\frac{\partial^2 D}{\partial\phi_a\partial\phi_b}\right) = -\frac{1}{2}\delta_{ab}\left[D^{-1}_{a+1,a}\mathrm{e}^{i\phi_a+\mu} - D^{-1}_{a,a+1}\mathrm{e}^{-i\phi_a-\mu}\right] \tag{3.29}$$

and

$$\mathrm{tr}\left(D^{-1}\frac{\partial D}{\partial\phi_a}D^{-1}\frac{\partial D}{\partial\phi_b}\right) =$$

$$(\frac{i}{2})^2\Big\{\mathrm{e}^{i\phi_a+\mu}\mathrm{e}^{i\phi_b+\mu}(D^{-1})_{b+1,a}(D^{-1})_{a+1,b} + \mathrm{e}^{i\phi_a-\phi_b}(D^{-1})_{b,a}(D^{-1})_{a+1,b+1} +$$
$$\mathrm{e}^{-i\phi_a+\phi_b}(D^{-1})_{a,b}(D^{-1})_{b+1,a+1} + \mathrm{e}^{-i\phi_a-\mu}\mathrm{e}^{-i\phi_b-\mu}(D^{-1})_{b,a+1}(D^{-1})_{a,b+1}\Big\}$$

With these ingredients it is now possible to compute the hessian and find the directions tangent to the thimble attached to the solution we have found in the constant field subspace. The tangent space for a typical set of parameters is given in Table (3.1).

$$\begin{bmatrix} 0.810 \\ -0.493 \\ -0.316 \\ 0.0 \end{bmatrix} \quad \begin{bmatrix} -0.288 \\ -0.288 \\ -0.288 \\ 0.866 \end{bmatrix} \quad \begin{bmatrix} -0.102 \\ -0.650 \\ 0.752 \\ 0.0 \end{bmatrix} \quad \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}$$

Table 3.1: Listed is a basis of the tangent space of the thimble at the extrema $\phi_t = i\phi^*$ for the parameters $N_t = 4,\ \hat{m} = 1.0,\ \hat{g} = 1.0,\ \hat{\mu} = 0.1$.

It can be noticed from Table (3.1) that the tangent vectors are purely real, so the tangent plane is parallel to the real plane. So far there is no proof that the tangent plane is always parallel to the real axis in the Thirring Model, however this feature holds for every combination of parameters tried in this work. The fact that the tangent plane is parallel to the original domain of integration is computationally useful. For one, it is not necessary to store tangent vectors for simulations, which is helpful for large lattices where storing the tangent plane might require memory not available. Additionally, during a Monte Carlo simulation, where one needs to constantly map from the coordinate manifold to the deformed manifold, the map is trivial: one simply adds $i\phi^*$ to a point on the coordinate manifold to compute the corresponding point on the deformed manifold.

3.4 The Flow

Recall that our strategy for taming the sign problem for the Thirring model to first shift the manifold, then flow the shifted manifold. We have discussed the details of the shift in the previous section and we now move on to the flow.

The exact form for a flowed point is

$$\tilde{\phi}(\phi) = \phi + \int_0^{T_{\text{flow}}} dt \left(\frac{\partial S(\phi(t))}{\partial \phi} \right)^* . \tag{3.30}$$

The derivative of the action is

$$\frac{\partial S}{\partial \phi_t} = \frac{\sin \phi_t}{\hat{g}^2} + \frac{i}{2}\left(D^{-1}_{t,t+1} \mathrm{e}^{i\phi_t + \hat{\mu}} + D^{-1}_{t+1,t} \mathrm{e}^{-i\phi_t - \hat{\mu}} \right) \tag{3.31}$$

and in our simulations we solve Eq. (3.30) numerically using a Cash-Karp integrator, which is an adaptive step size integrator of order $\mathcal{O}(\Delta t^5)$ [103]. Such an adaptive integrator is helpful for efficiently handling the right hand side of Eq. (3.30) which grows rapidly as the flow time is increased. Similarly the exact form of the Jacobian is given by

$$J(\phi) = \mathbb{1} + \int_0^{T} dt \, \Big(H(\phi(t)) J(\phi(t)) \Big)^* \tag{3.32}$$

which is evolved with the same numerical method that flows points. The explicit form of the Hessian is given in the last section.

We conclude by noting that, if a point where $\det D(\phi) = 0$ is approached along a flow trajectory, the holomorphic gradient flow causes the flow trajectory to evolve very quickly to points of high action. In these situations, the adaptive step size of the Cash-Karp integrator tends to shrink without bound. If this happens, then the Markov chain gets stuck at a particular configuration rendering the calculation useless.

In our simulations, if the step size shrinks past a pre-specified minimum, the point is rejected automatically and a new point is drawn to continue the Metropolis chain. This procedure prevents the Metropolis chain from getting stuck to a particular configuration and introduces at worst exponentially small biases which cannot be resolved by a Monte Carlo of reasonable length. This is because the action monotonically grows along a flow trajectory, so if a $\det D(\phi) = 0$ point is approached, the action will shoot up to very large values and never decrease. Therefore, since the probability of a configuration is $\mathrm{e}^{-\mathrm{Re}\, S}$, points whose trajectory approach zeros of the determinant give vanishingly small contributions to the path integral and can be rejected without consequence.

3.5 Results

The (0+1)d Thirring model is solvable, so we will calculate observables both exactly and on the lattice; the agreement between these two approaches validates the holomorphic gradient flow approach. The results in this section are from [18].

To begin, the partition function can be computed exactly using techniques from [104, 105]. One finds:

$$Z = \frac{\mathrm{e}^{-N\alpha}}{2^{N-1}} \left[I_1^N(\alpha) \cosh(N\hat{\mu}) + I_0^N(\alpha) \cosh\!\big(N \sinh^{-1}(\hat{m})\big) \right] \tag{3.33}$$

where $\alpha = \frac{1}{2\hat{g}^2}$ and I_n are modified Bessel functions of the first kind, which are defined by the integral equation [106]

$$I_n(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi\ \mathrm{e}^{z\cos(\phi)}\mathrm{e}^{in\phi}\ . \tag{3.34}$$

The charge density and chiral condensate are given by derivatives of the partition function:

$$\begin{aligned}\langle n\rangle &\equiv \frac{1}{\beta}\frac{\partial}{\partial\mu}\log Z = \frac{1}{N}\frac{\partial}{\partial\hat{\mu}}\log Z \\ &= \frac{I_1^N(\alpha)\sinh(N\hat{\mu})}{I_1^N(\alpha)\cosh(N\hat{\mu}) + I_0^N(\alpha)\cosh\bigl(N\sinh^{-1}(\hat{m})\bigr)}\end{aligned} \tag{3.35}$$

$$\begin{aligned}\langle\overline{\psi}\psi\rangle &\equiv \frac{1}{\beta}\frac{\partial}{\partial m}\log Z = \frac{1}{N}\frac{\partial}{\partial\hat{m}}\log Z \\ &= \frac{(1+m^2)^{-1/2}I_0^N(\alpha)\sinh\bigl(N\sinh^{-1}(\hat{m})\bigr)}{I_1^N(\alpha)\cosh(N\hat{\mu}) + I_0^N(\alpha)\cosh\bigl(N\sinh^{-1}(\hat{m})\bigr)}\ .\end{aligned} \tag{3.36}$$

Let us begin by looking at the average phase $\langle \mathrm{e}^{-iS_I}\rangle$, which is a diagnostic measure of the severity of the sign problem. When the average phase has close to unit modulus, this means that the phase hardly fluctuates and the sign problem is gentle. On the other hand when the average phase has close to zero modulus, the phase fluctuates a lot and therefore the sign problem is severe. It can be seen in Fig. (3.3) that the sign problem becomes severe at high densities. We find that the severity of the sign problem rapidly increases when the lattice begins to become populated with charge. This happens when $\mu/m_f \simeq 1$ where m_f is the physical mass of the fermion field ψ. Therefore, for $\mu > m_f$ it is necessary to deform the domain of integration. Let us now discuss this process.

It is helpful to see the flow happen directly. In particular, the connection between flowed manifolds and thimbles becomes particularly obvious when one sees a manifold deform into a set of thimbles. Unfortunately the deformation

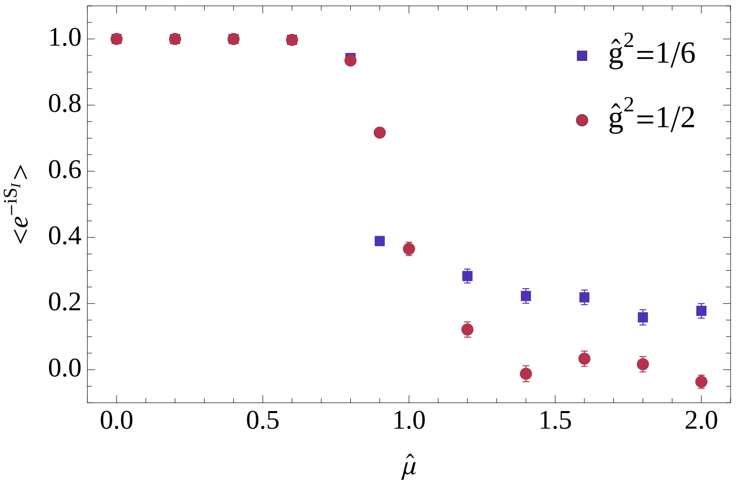

Figure 3.3: The average phase on the original domain of integration S_1^N for $N = 8,\ \hat{m} = 1$ and $\hat{g}^2 = 1/6$ (blue) and $\hat{g}^2 = 1/2$ (red). Notice the rapid decay of the average phase as the density of the system is increased.

occurs in a high dimensional complex space and so visualization is not easy. There is one device we have developed to aid visualization however, and this is called the "constant field subspace". The constant field subspace is the set of all fields which are constant in spacetime $\{\phi = (\phi_1, ..., \phi_N) \mid \phi_t = \phi\ \forall t\}$. Notice that this is a one real dimensional subspace of the original domain of integration. For any action with translation invariance, any field in the constant field subspace of the original domain of integration will remain in the constant field subspace under the holomorphic gradient flow. This is simple to show: by translation invariance every component of $\frac{\partial S}{\partial \phi_t}$ is the same when evaluated at a point which is constant in spacetime. Since all components of the gradient are equal, and points move in the direction of the gradient of the action, the evolution of a point starting in the constant field subspace is the same in all directions. This means that if $\phi(t = 0)$ is constant in spacetime, then the flowed point $\tilde{\phi} = \phi(t)$ will be constant in spacetime too. This means that one dimension of a flowed surface can be plotted. A plot of the constant field subspace of a sequence of flowed manifolds is shown

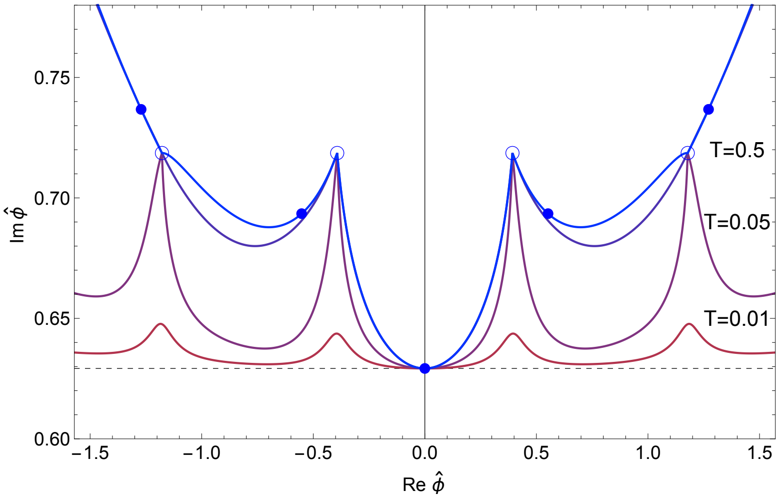

Figure 3.4: Here is the evolution of the constant field subspace under the holomorphic gradient flow. Before flowing, the domain of integration is shifted to the constant field critical point $\phi = (i\phi^*, ..., i\phi^*)$. This initial manifold, which is flat and parallel to the real axis, has a direction along the constant field subspace which is plotted as the dotted gray line. These points are then subjected to various amounts of flow. The red curve shows the result of flowing all points in the constant field subspace for $T = 0.01$. The purple and blue curves are the manifolds obtained by flowing for $T = 0.05$ and $T = 0.5$ respectively. **Note:** The solid blue points denote extrema of the action and the hollow blue points denote zeros of the fermion determinant.

in figure Fig. (3.4). Looking at Fig. (3.4) helps one to see the flow generating a family of manifolds which interpolates between the shifted real plane and a set of thimbles. The blue curves in Fig. (3.4) are constant field subspace projections of the thimbles which contribute to the path integral. The red then purple manifolds clearly converge to the blue thimbles.

It is also instructive to examine the action along flowed manifolds. In Fig. (3.5) we have plotted the action on manifolds flowed by various amounts. For visualization purposes, we choose to represent each point on a flowed manifold by its projection on the real axis. That is, if $\mathcal{M}_T$ is a flowed manifold, and if $\tilde{\phi} \in \mathcal{M}_T$, then in our plot the point $\tilde{\phi}$ corresponds to Re$\tilde{\phi}$. The first thing to notice is that paths of stationary phase indeed emerge by flowing the shifted manifold. This is

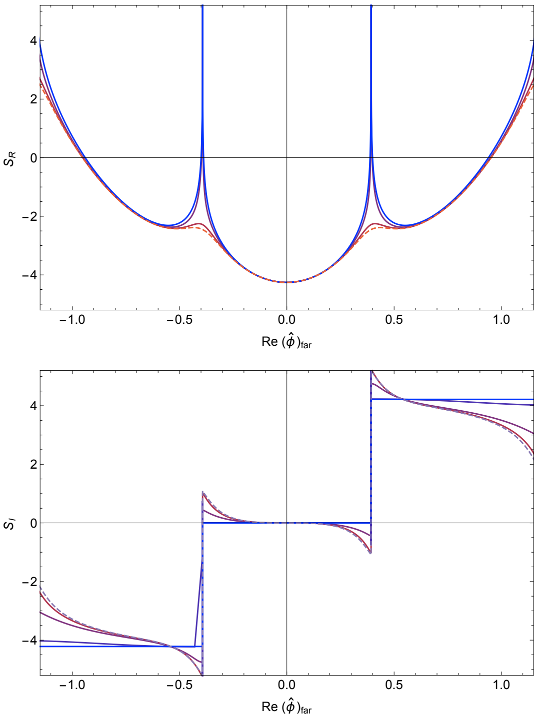

Figure 3.5: Here we plot the real (top) and imaginary (bottom) parts of the action, as a function of the horizontal projection of the point on the manifold. In the top figure, the dotted red line is the real part of the action along the shifted but unflowed manifold. The ever bluer profiles are obtained from manifolds flowed by $T = 0.01,\ 0.05,\ 0.50$. From viewing the top figure it is evident that a large amount of flow creates action barriers between thimble, and the bottom figure shows that the manifolds obtained have stationary phase.

demonstrated by the convergence of S_I to a piecewise constant function. Evidently three thimbles intersect the constant field subspace in this model. Furthermore, from the top plot one finds that large flow times generate action barriers between regions of parameter space.

A close examination of the behavior of the action as a function of flow time demonstrates how the holomorphic gradient flow tames sign problems. First, along a flow trajectory, which takes a point on the coordinate manifold to a point on a flowed manifold, the real part of the action strictly increases while the imaginary part of the action is constant. This is easily derived from the definition of the holomorphic gradient flow:

$$\frac{d}{dt}S(\phi(t)) = \frac{\partial S(\phi(t))}{\partial \phi}\frac{d\phi(t)}{dt} = \frac{\partial S(\phi(t))}{\partial \phi}\Big(\frac{\partial S(\phi(t))}{\partial \phi}\Big)^* = \left|\frac{\partial S(\phi(t))}{\partial \phi}\right|^2 > 0 \quad (3.37)$$

A consequence of this observation is that, the regions of support of the probability distribution $p(\phi) = \mathrm{e}^{-S_R(\tilde{\phi}(\phi))}$ on the parameterization manifold monotonically shrink as a function of flow time. This is because as the action of a point increases its probability decreases. In fact, a point ϕ in the parameterization manifold typically achieves a very low probability in a finite flow time due to the ever increasing action caused by the holomorphic gradient flow. For a large flow time, the only points on the parameterization manifold with any reasonable chance of being sampled are those points which are in a small neighborhood of those few points which flow to extrema of the action. This is because the only points which ever stop moving are those for which $\frac{\partial S}{\partial \phi} = 0$, which are those points which asymptote to an extremum of the action. For a finite flow time, there will be a set of points around a point which flows to an extremum which have small enough actions that they contribute significantly to the path integral.

Now, in addition to the real part of the action ever increasing, the imaginary part of the action stays constant along a flow trajectory. Since the integrand of the path integral is a holomorphic function and is therefore continuous, small neighborhoods around any given point have mildly fluctuating phases. The fluctuations of the phase in field space are not affected by the holomorphic gradient flow, and therefore the small pockets of parameterization space which have finite probability after a large amount of flow time have stationary phase. This combination of effects is how the holomorphic gradient flow works: applying the holomorphic gradient flow results in a probability distribution on the parameterization manifold that has pockets of high probability and stationary phase.

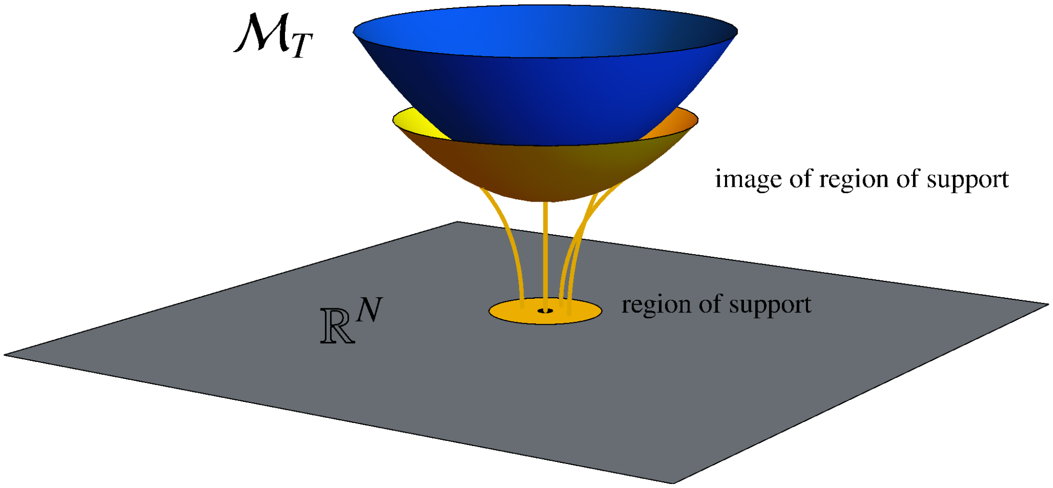

Figure 3.6: Here we depict schematically how regions of parameter space are rendered important and otherwise by the holomorphic gradient flow. The black point is a point which flows to an extremum of the action. For a given finite flow time, there is a region of support around the this black point which has appreciable probability, and all other points are mapped to high actions. The yellow region on $\mathbb{R}^N$ is the small neighborhood which has high probability. Its image under the flow is the yellow conic shape, with particular flow trajectories shown in the yellow lines. Away from this neighborhood (the gray) points are mapped to high action.

Now we must point out an important detail to the argument above. The argument relied on the fact that along a flow trajectory the probability

$$p(\phi) \propto \mathrm{e}^{-\mathrm{Re}\, S(\tilde{\phi}(\phi))} \tag{3.38}$$

of a point ϕ on the parameterization manifold strictly decreases with flow. However, the integral includes a jacobian. One approach is to view the jacobian as an extra factor by which every observable must be reweighted. That is, we write observables as

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O}\ \det J \rangle_{S_R}}{\langle \det\ J \rangle_{S_R}} \tag{3.39}$$

In this case the previous discussion is unchanged: one samples with respect to $p(\phi) \propto \mathrm{e}^{-\mathrm{Re}\, S(\tilde{\phi}(\phi))}$ and flow creates a probability distribution which becomes strictly increasingly peaked as a function of flow. The only detail is that observables are reweighted with the jacobian.

In our simulations, however, we sample with respect to

$$p(\phi) \propto \mathrm{e}^{-\mathrm{Re}\ S_{\mathrm{eff}}(\phi)},\ S_{\mathrm{eff}}(\phi) = S(\tilde{\phi}(\phi)) - \ln\ \det J(\phi) \tag{3.40}$$

which includes the jacobian. Under the holomorphic gradient flow, the effective action evolves as

$$\frac{d}{dt} S_{\mathrm{eff}}(\phi(t)) = \left| \frac{\partial S(\phi(t))}{\partial \phi} \right|^2 - \mathrm{tr}\big(J^* J^{-1} H^*\big)\ . \tag{3.41}$$

which is complicated by the $\mathrm{tr}(J^* J^{-1} H^*)$ term. While it is difficult to describe the analytic behavior of this probability distribution, in practice we find that $\mathrm{e}^{-\mathrm{Re}\ S_{\mathrm{eff}}(\phi)}$ behaves very similarly to $\mathrm{e}^{-\mathrm{Re}\ S(\phi)}$: flowing the manifold generates pockets of high probability and stationary phase on the parameterization manifold.

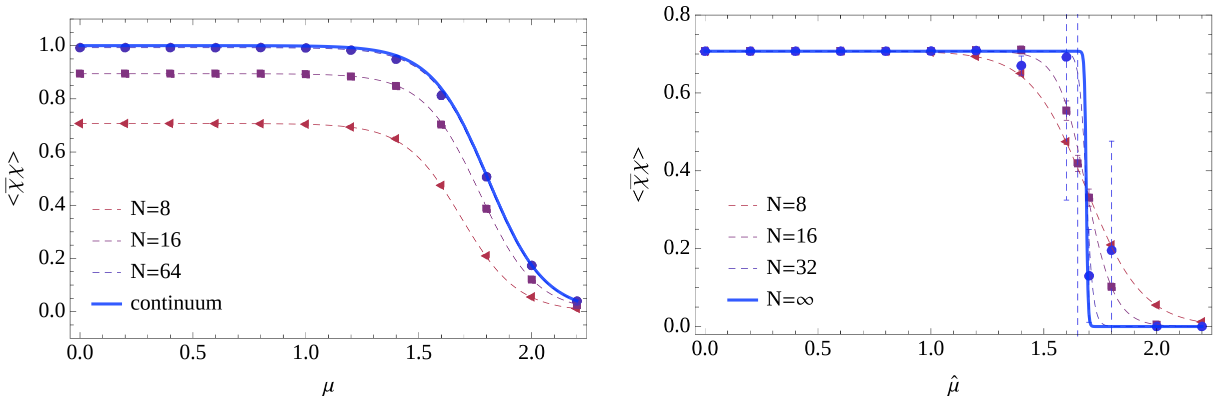

Figure 3.7: Here we show the chiral condensate computed in the continuum (left) and zero-temperature (right) limits. The $N = 8$ calculations are have parameters set at $\hat{m} = 1.0,\ \hat{g}^2 = 1/2$, which is a strongly coupled case. In the left plot where the continuum limit is taken, red is a coarse lattice spacing and blue is a fine lattice spacing. In the right hand plot red is hot and blue is cold. The scaling of the parameters as the continuum and zero-temperature limits

Having demonstrated that there is a sign problem in Fig. (3.3), and having described how and why the holomorphic gradient flow works to tame it, let us finally examine some calculations on flowed manifolds.

In Fig. (3.7) we compute the chiral condensate $\langle\overline{\psi}\psi\rangle$ as a function of the chemical potential $\hat{\mu}$. We explore both the continuum limit and the zero temperature limit. We expect the sign problem to worsen as a function of the spacetime volume, and we indeed find this to be the case. In the left hand panel of Fig. (3.7) we take the continuum limit of the theory at fixed lattice volume (i.e. fixed temperature). The continuum limit is taken by taking $N \to \infty$ while keeping $\hat{m}N, \hat{g}^2N, \hat{\mu}N$ fixed [18]. We find that a simple shift to the tangent plane suffices to tame the sign problem enough to compute observables for an entire sweep of $\hat{\mu}$. No flow is required for these calculations.

On the other hand, as we decrease the temperature the sign problem worsens. This is evidenced by the right hand plot of Fig. (3.7) where the error bars grow as the zero temperature limit is taken. The zero temperature limit calculations were also done on the shifted tangent plane, and it is clear that the tangent plane is good enough for certain temperatures but not others. The bluest data in Fig. (3.7) for example has a bad enough sign problem that at the transition $\hat{\mu}$, which occurs roughly between $1.6 \leq \hat{\mu} \leq 1.7$, the value of the condensate is completely unknown due to large statistical uncertainties.

There are two lessons here worth pointing out. The first lesson is that often it is sufficient to simply shift the manifold of integration to tame the sign problem enough to do calculations. This holds for a variety of models, and this is useful because there is almost zero cost to computed on a shifted manifold. The second lesson we find is that the sign problem becomes severe when thermodynamic observables sharply change. This holds for a variety of models as we will see in later chapters.

N	$\langle\bar{\chi}\chi\rangle$	$\langle\bar{\chi}\chi\rangle_{exact}$
16	0.330(13)	0.353
32	0.345(46)	0.353
64	0.375(48)	0.353

Table 3.2: Value of the condensate in the low temperature ($N \to \infty$) limit and $\mu = 1.688$ obtained with flow time $T_{flow} = 0.5$. These results and its error bars should be compared to the ones in the right panel of Fig. (3.7) obtained with $T_{flow} = 0$.

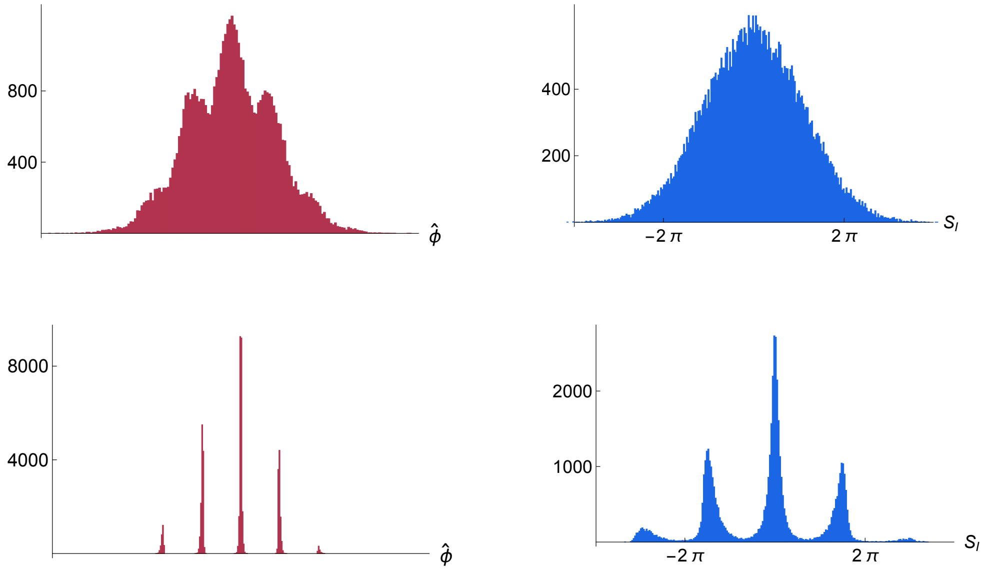

Figure 3.8: Here we show the results of a Monte Carlo on $T_{flow} = 0.0$ (top) and $T_{flow} = 0.5$ (bottom) manifolds at $N = 32$. The left hand plots are a probability distribution of the fields on the constant field subspace of the parameterization manifold, and the right hand plots are a distribution of the imaginary part of the action. The diffusiveness of the imaginary part of the action on the top row is responsible for the sign problem. It is evident from the bottom row that flowing the manifold creates small pockets in the parameterization manifold with high probability (left) and relatively stationary phase (right).

Let us now consider the transition region in Fig. (3.7), where a simple tangent shift is not sufficient to tame the sign problem. Since the chemical potential is basically a step function at low temperatures, a precise chemical potential of $\hat{\mu} = 1.688$, which lies at the middle of the jump, is chosen as a representative point to study.

We first note that the flow time must be tuned to a certain extent. On the

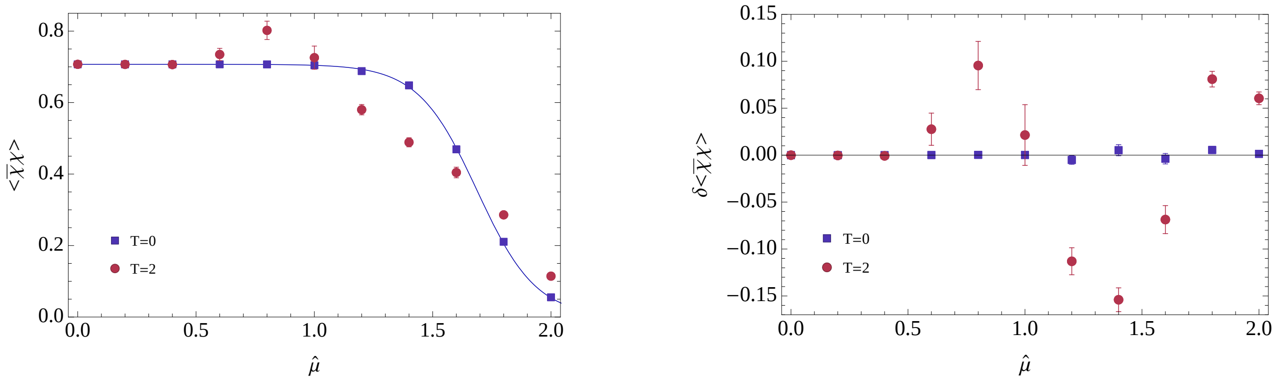

Figure 3.9: Here we show the condensate (left) and the deviation of the condensate from the exact result (right) as a function of the chemical potential $\hat{\mu}$. The parameters for these calculations are $N = 8$, $\hat{m} = 1.0$ and $\hat{g} = 1/2$. Under these conditions the sign problem is non-negligible, but can be handled with high statistics. The blue points are results from $T_{flow} = 0.0$ calculations while the red points are obtained fro $T_{flow} = 2.0$ calculations.

one hand, a large enough flow time will generate a set of thimbles upon which the phase is piecewise constant. This is good because this tames the sign problem. On the other hand, because the holomorphic gradient flow creates pockets with high probability, too much flow can trap a Monte Carlo of reasonable length to a single thimble which can lead to systematic errors in observables. We will see this later on. For the $\hat{\mu} = 1.688$ point, after some trail and error we find a flow time of $T_{\text{flow}} = 0.5$ to be sufficient to tame the sign problem without being trapped to a single thimble.

Calculating on these flowed manifolds, we find the results Table (3.2). Comparing these results to those found in Fig. (3.7), it is clear that the flowed manifolds have a much tamer sign problem than the tangent plane.

For visual purposes we show histograms of the action of the field configurations obtained on flowed and not-flowed manifolds in Fig. (3.8). The top row is obtained on the tangent plane while the bottom row is obtained from a $T_{flow} = 0.5$ manifold. It is clear that the Monte Carlo samples paramter space more smoothly on the tangent plane, and this results in a more diffuse distribution of S_I. On the other hand, as can be seen from the bottom row of Fig. (3.8), after applying the holomorphic gradient flow the important pockets of the parameterization manifold are isolated and the imaginary part of the action fluctuates less. This tames the sign problem.

We end this section with a caution: don't get trapped. As has been stated previously, a large amount of flow can trap a simulation to a single thimble and this leads to systematic uncertainties in observables. We demonstrate that this happens in Fig. (3.9). While the $T_{flow} = 0.0$ results match with the exact result, one sees significant deviation from the exact result at $T_{flow} = 2.0$. This is because the $T_{flow} = 2.0$ result is obtained by integrating over a single thimble, while many thimbles contribute in this model.

Chapter 4

Optimizations for Simulations

While in (0+1) dimensions it is relatively simple to calculate on flowed manifolds, higher dimensions are more difficult due to the increased number of degrees of freedom. In this chapter we will discuss two optimizations we have developed to alleviate this difficulty [21,94]. One is a particular scaling of Metropolis proposals and the other is a cheap way to compute approximate jacobians.

4.1 Optimizing Metropolis Proposals

We use a metropolis algorithm to sample fields on flowed manifolds. Since a metropolis style Monte Carlo sampling relies on proposing small changes $\phi \to \phi'$ on the coordinate manifold, and then accepting this proposal with probability

$$\Pr = \min\{1, \mathrm{e}^{-(S(\phi')-S(\phi))}\} \tag{4.1}$$

it is necessary that changes in the action $S(\phi') - S(\phi)$ due to the proposal are not too large. One typically tries to make the change in the action due to a proposal to be $\mathcal{O}(1)$, which achieves an acceptance rate of roughly 50%. Such an acceptance rate allows the Monte Carlo chain to sample the configuration space without being stuck to a particular configuration for a long time.

To achieve a reasonable acceptance rate on a flowed manifold can be hard however, because a small displacement in coordinate space can be mapped under the flow to a large displacement on a flowed manifold. Even for a free theory, a point on the coordinate manifold ϕ flows as $\tilde{\phi}(t) = \sum_i \phi_i(0)\mathrm{e}^{\lambda_i t}\rho_i$, so small changes in the $\phi_i(0)$ are exponentially magnified.

We account for this rapid flowing of points to achieve a reasonable acceptance rate by scaling the width of the proposal distribution in different directions. Before describing how we scale proposals, let us understand why this might be necessary. First, any point on the coordinate manifold (which in the case of the (1+1) thirring model is the tangent plane of the thimble in the constant field subspace) is of the form $\phi = \sum c_i\rho_i$ where c_i are real and ρ_i are the Takagi vectors associated with the extremum of the action. In our simulations, we make proposals by changing the $\{c_i\}$ by small amounts. Near the extremum, the change in the action due to the proposal $c_i \to c_i + \delta c_i$ is

$$\delta S = \epsilon \sum_i 2c_i\delta c_i\lambda_i\mathrm{e}^{2\lambda_i T} \ . \tag{4.2}$$

for a flow time of T. One sees that due to the eigenvalues λ_i, different directions in the tangent plane contribute different amounts to the change in action. Clearly it is sensible to propose smaller changes in directions with large eigenvalues. Also one sees that proposals need to be smaller for larger flow times to keep the change in action $\mathcal{O}(1)$.

We therefore propose as follows:

$$c_i \to c_i + \epsilon \times \left(\frac{\mathrm{e}^{-\lambda_i T}}{\sqrt{\lambda_i}}\right) \tag{4.3}$$

where ϵ is a uniform random variable on the interval $[-\Delta, \Delta]$ where Δ is tuned as desired. The factor of the $1/\sqrt{\lambda_i}$ absorbs some of the λ_i dependence in each direction and the factor of $\mathrm{e}^{-\lambda_i T}$ absorbs some of the flow time dependence. These scalings are extremely important. Without them, even simulations with relatively small flow times have very low acceptance rates. With the scalings in place, the single proposal parameter to tune, Δ, varies little from simulation to simulation.

4.2 Jacobian Estimators

With a cost scaling as $\mathcal{O}(N^8 \times N_{step})$, computing the Jacobian in a theory with fermions is the most expensive part of a calculation using the holomorphic gradient flow. With a small number of degrees of freedom this cost can be dealt with by brute force, however brute force quickly becomes ineffective as the lattices become larger. To deal with this issue we have developed several "estimators", quantites which track the behavior of the jacobian but which are less expensive to compute [94] . Recall that any observable can be expressed as

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O} \mathrm{e}^{-i \mathrm{Im} S_{eff}} \rangle_{\mathrm{Re}\ S_{eff}}}{\langle \mathrm{e}^{-i \mathrm{Im} S_{eff}} \rangle_{\mathrm{Re}\ S_{eff}}} \tag{4.4}$$

where $S_{eff}(\phi) = S(\tilde{\phi}(\phi)) - \ln \det J(\phi)$ and $\langle\ \cdot\ \rangle_{\mathrm{Re}\ S_{eff}}$ means to average with respect to the probability distribution

$$p(\phi) = \frac{e^{-\mathrm{Re}\ S_{eff}(\phi)}}{\int_{\mathbb{R}^N} D\phi\ \mathrm{e}^{-\mathrm{Re} S_{eff}(\tilde{\phi}(\phi))}} \tag{4.5}$$

The idea behind an estimator is to sample with respect to a different probability distribution, in particular one that is easier to sample from, then reweight the difference [94]. If we call $\tilde{S}_{eff}(\phi) = S(\tilde{\phi}(\phi)) - \ln \det \tilde{J}(\phi)$ the action defined with respect to an estimator jacobian $\tilde{J}$, then any observable can be re-written *exactly* as

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O} \mathrm{e}^{-i \mathrm{Im} \tilde{S}_{eff}} \mathrm{e}^{\Delta \ln\ \det J} \rangle_{\mathrm{Re}\ \tilde{S}_{eff}}}{\langle \mathrm{e}^{-i \mathrm{Im} \tilde{S}_{eff}} \mathrm{e}^{\Delta \ln\ \det J} \rangle_{\mathrm{Re}\ \tilde{S}_{eff}}} \tag{4.6}$$

where $\langle \ \cdot \ \rangle_{\mathrm{Re}\ \tilde{S}_{eff}}$ means to now average with respect to the new probability distribution

$$\tilde{p}(\phi) = \frac{e^{-\mathrm{Re}\ \tilde{S}_{eff}(\phi)}}{\int_{\mathbb{R}^N} D\phi\ \mathrm{e}^{-\mathrm{Re}\tilde{S}_{eff}(\tilde{\phi}(\phi))}} \tag{4.7}$$

and $\mathrm{e}^{\Delta \ln\ \det J} = \mathrm{e}^{\ln\ \det J}/\mathrm{e}^{\ln\ \det \tilde{J}}$ is a complex number with both phase and magnitude which we will call the "reweighting fator". Eq. (4.6) tells us that one can obtain correct results by sampling from any probability distribution desired provided the difference between the sampled distribution $\tilde{p}$ and the actual distribution p is accounted for by reweighting observables with the appropriate reweighting factor. While Eq. (4.6) is an exact relation, its usefulness hinges on $\tilde{J}$ being close enough to J.

A key point for the utility of an estimator is that at each step of the Monte Carlo chain only the cheaper $\tilde{J}$ is needed. One updates the Markov chain with the estimator, which is a cheap procedure, and take measurements every so often when the Markov chain decorrelates. Only on measurement steps does one compute the reweighting factor. This calculating the reweighting factor only on the small subset of configurations where measurements are taken is what allows the speedup to happen.

We have developed two estimators for use in simulations. One is a good approximation when the jacobian is nearly real and the other is a good approximation when the action is nearly quadratic. The first estimator is

$$\ln\ \det \tilde{J}_1(t) = \int_0^T dt\ \ \mathrm{tr}\, H^*(t)\ , \tag{4.8}$$

where T is the flow time. This estimator is the log det of the solution to the differential equation:

$$\frac{d}{dt}\tilde{J}_1 = \big(H\big)^* \tilde{J}_1 \tag{4.9}$$

which differs from the actual jacobian evolution

$$\frac{d}{dt}J = \big(HJ\big)^* \tag{4.10}$$

by only a complex conjugate. Indeed the solution of Eq. (4.9) is

$$\tilde{J}_1(t) = \tilde{J}_1(0)\; P\; \exp\Big(\int_0^t dt'\;\, H^*(t')\Big) \tag{4.11}$$

where P denotes path ordering along a flow trajectory. Initializing $\tilde{J}_1(0)$ to be a matrix of unit determinant, Eq. (4.8) follows. We see that the estimator tracks the jacobian well when the jacobian is nearly real. This is the approximator that we use in the (1+1) Thirring model, which will be discussed in the next chapter.

To establish the cost of the estimator, note that the cost of computing Eq. (4.8) is the number of time steps N_{step} times the cost of each step. In the case of fermions, as was pointed out in Sec. 2.1 the cost of computing an element of the hessian $H_{xy} \sim D^{-1}_{xy} D^{-1}_{xy}$ is the cost of inverting the fermion matrix which is $\mathcal{O}(N^3)$ where N is the number of lattice points. Computing the trace costs another factor of N and therefore the estimator Eq. (4.8) costs $\mathcal{O}(N^4 \times N_{\text{step}})$, which is a substantial improvement over the $\mathcal{O}(N^8 \times N_{\text{step}})$ cost of the actual jacobian.

The second estimator we use is

$$\ln\, \det\tilde{J}_2(t) = \int_0^T dt\; \sum_a \rho^a \big(H(t)\rho^a\big)^*\;, \tag{4.12}$$

where $\{\rho^a\}$ is an orthonormal basis of tangent vectors to the thimble at the extremum in the constant field subspace (which recall in the case of the finite density Thirring Model is just the standard basis of $\mathbb{R}^N$). To motivate this estimator, note that in the vicinity of the extremum where the action is nearly quadratic, this is an exact solution. To see this, notice that $\frac{d}{dt}J = \big(HJ\big)^*$ has the exact solution:

$$J(t) = J(0) + \int_0^t dt_1 \; H^*(t_1)J^*(0) + \int_0^t \int_0^{t_1} dt_1 dt_2 \; H^*(t_1)H(t_2)J(0) + ... \tag{4.13}$$

Near the extremum the Hessian is constant, in which case

$$J(t) = J(0) + tH^*J^*(0) + \frac{t^2}{2}H^*HJ(0) + ... \tag{4.14}$$

Using the freedom one has to initialize the Jacobian to any matrix desired, so we choose $J(0) = \big(\rho^1, \rho^2, ..., \rho^N\big)$ to be a row vector of Takagi vectors. Using the fact that $\big(H\rho^a\big)^* = \lambda_a \rho^a$ one finds

$$\big(HJ(0)\big)^* = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_N)J(0) \equiv \Lambda J(0) \tag{4.15}$$

This allows us to solve

$$\begin{aligned} J(t) &= J(0) + tH^*J^*(0) + \frac{t^2}{2}H^*HJ(0) + ... \\ &= \big(\mathbb{1} + t\Lambda + \frac{(t\Lambda)^2}{2} + ...\big)J(0) \\ &= \exp(t\Lambda)J(0) \end{aligned} \tag{4.16}$$

in which case

$$\ln \det J(t) = t\sum_a \lambda_a = \int_0^T dt \; \sum_a \rho^a \big(H(t)\rho^a\big)^* \; . \tag{4.17}$$

So we see that Eq. (4.12) is a good estimator to the extent that the action is quadratic. The cost of computing Eq. (4.12) is $\mathcal{O}(N^6 \times N_{\text{step}})$. Since:

1. $H_{xy} \sim D^{-1}_{xy} D^{-1}_{xy}$ costs N^3

2. For a single a, $\rho_x^a\big(H_{xy}\rho_y^a\big)^*$ costs N^2

3. The sum over a costs N

4. The previous steps must be repeated N_{step} number of times.

One sees that Eq. (4.12) is a more expensive estimator that Eq. (4.8), however Eq. (4.12) was still found to be useful as it tracks the real jacobian more closely than Eq. (4.8).

Before quantifying the utility of these estimators as we will in the following section, let us briefly repeat the analysis above for bosonic models. First, computing the actual jacobian evolution requires the evolution of $\frac{d}{dt}J = \big(HJ\big)^*$. For a bosonic theory, the Hessian is a sparse matrix and the right hand side of this equation can be written down exactly without having to perform a matrix multiplication. For example if we take massless ϕ^4 theory, an action is

$$S = \sum_{x\mu} \frac{(\phi_{x+\mu} - \phi_x)^2}{2} \tag{4.18}$$

where x ranges over a d dimensional hypercubic lattice and $\mu = 1, ..., d$. This action gives a sparse Hessian

$$H_{xy} = d\delta_{xy} - \frac{1}{2}\sum_{\mu} \delta_{y,x+\mu} + \delta_{y,x-\mu} \tag{4.19}$$

Consequently, each time step of $\frac{d}{dt}J = \big(HJ\big)^*$ costs only $\mathcal{O}(N^2)$ to compute. At the end of a flow trajectory a determinant must be computed, which costs $\mathcal{O}(N^3)$. Therefore the cost of computing ln det$J$ for a bosonic theory is either $\mathcal{O}(N^2 \times N_{\text{step}})$ or $\mathcal{O}(N^3)$; typically in our calculations $N >> N_{\text{step}}$ and the cost of computing ln det$J$ scales as $\mathcal{O}(N^3)$.

Now, for the first estimator

$$\ln\, \det \tilde{J}_1(t) = \int_0^T dt\ \ \mathrm{tr}\, H^*(t) \tag{4.20}$$

computing the trace for bosons costs N for each step, and so $\ln\, \det \tilde{J}_1(t)$ costs $\mathcal{O}(N \times N_{\text{step}})$. For the second estimator,

$$\ln\, \det \tilde{J}_2(t) = \int_0^T dt\ \sum_a \rho^a \big(H(t)\rho^a\big)^* \tag{4.21}$$

due to the sparsity of H, for fixed a, computing $\rho^a \big(H(t)\rho^a\big)^*$ costs $\mathcal{O}(N)$ and summing over a costs another factor of N. Therefore this estimator costs $\mathcal{O}(N^2 \times N_{\text{step}})$, which is less than the actual jacobain because $N_{\text{step}} << N$.

4.3 Results

We now test these estimators on the (0+1) Thirring model to quantitatively understand their performance. The results in this section come from [94].

A quantity which measures the quality of an estimator is the "statistical power", which is defined as follows. First, for a set of field configurations $\{\phi_1, \phi_2, ..., \phi_{\mathcal{N}}\}$ drawn from the distribution $\tilde{p}(\phi)$ define the quantity

$$w = \frac{\mathrm{e}^{\mathrm{Re}\ (\ln\, \det J - \ln\, \det \tilde{J}_i)}}{\sum_i \mathrm{e}^{\mathrm{Re}\ (\ln\, \det J - \ln\, \det \tilde{J}_i)}} \tag{4.22}$$

where the sum runs over the elements of the sample drawn. Up to an overall normalizaton, w is the factor by which each configuration is reweighted. The statistical power Σ is a function of the set of $\{w_1, w_2, ..., w_{\mathcal{N}}\}$ defined as

$$\Sigma = \frac{1}{\mathcal{N}} \frac{\langle w \rangle}{\langle w^2 \rangle}\ . \tag{4.23}$$

The statistical power is an observable which takes on values between $1/\mathcal{N} \leq \Sigma \leq 1$ which tells us what fraction of the sample is rendered statistically significant by the process of reweighting. Consider the case where $\mathrm{e}^{\mathrm{Re}\ (\ln\ \det J - \ln\ \det \tilde{J}_i)} = \mathrm{const}$, i.e. the difference between the estimator and the actual jacobian is constant. In this case our estimator is perfect: sampling with respect to an estimator which differs from the real jacobian by a constant is the same as sampling with respect to the jacobian. In fact, $w_i = 1/\mathcal{N}$ for every configuration and

$$\Sigma = \frac{1}{\mathcal{N}} \frac{\langle w \rangle}{\langle w^2 \rangle} = 1 \tag{4.24}$$

In other words, every single configuration is equally statistically significant. On the other extreme, consider a sample were $\mathrm{e}^{\mathrm{Re}\ (\ln\ \det J - \ln\ \det \tilde{J}_i)}$ is very large for a single configuration, say the first one, and much lower values for all other configurations. This is what would happen if the estimator were poor: the difference between the actual jacobian and the estimator fluctuates wildly over the sample. In this case

$$w_1 = \frac{\mathrm{e}^{\mathrm{Re}\ (\ln\ \det J - \ln\ \det \tilde{J}_i)}}{\langle \mathrm{e}^{\mathrm{Re}\ (\ln\ \det J - \ln\ \det \tilde{J}_i)} \rangle} = 1 \tag{4.25}$$

and $w_i \simeq 0$ for all others. In this case $\langle w \rangle = \langle w^2 \rangle = \frac{1}{\mathcal{N}}$ and so

$$\Sigma = \frac{1}{\mathcal{N}} \tag{4.26}$$

In other words, effectively only 1 out of $\mathcal{N}$ configurations contributes to the sample after reweighting.

We begin our analysis by noting that for the Thirring model (in any number of spacetime dimensions), the two estimators of the last section are the same. This

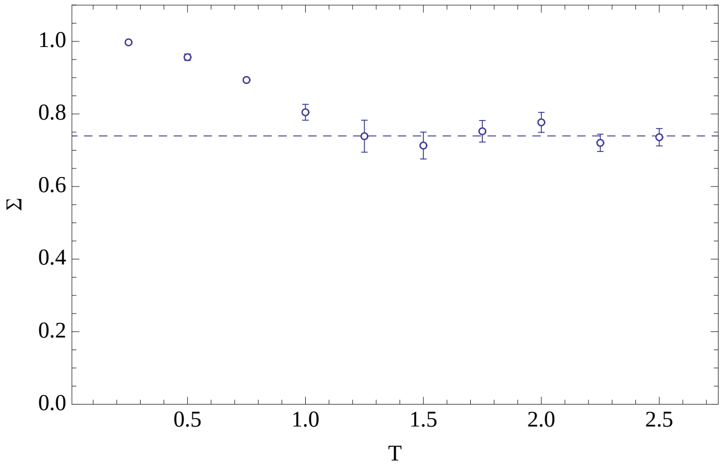

Figure 4.1: Here we plot the statistical power for the estimator $\ln\,\det\tilde{J}_1 = \ln\,\det\tilde{J}_2$ in the (0+1) Thirring model. The parameters for these calculations are $N = 8$, $\hat{g}^2 = 1/2$, $\hat{m} = 1.0$ and $\hat{\mu} = 1.0$. The dotted line is an average over the last six points. The convergence to $\Sigma \sim 0.75$ indicates that effectively 75% of the sampled configurations contribute to observables after the process of reweighting.

is because

$$\begin{aligned}\ln\,\det\tilde{J}_2(t) &= \int_0^T dt\;\sum_a \rho^a \big(H(t)\rho^a\big)^* = \int_0^T dt\;\operatorname{tr}\Big(H^*(t)\underbrace{\sum_a \rho^{a*}\otimes\rho^a}_{=\mathbb{1}}\Big) \qquad (4.27)\\ &= \int_0^T dt\;\operatorname{tr}\big(H^*(t)\big) = \ln\,\det\tilde{J}_1(t)\;,\end{aligned}$$

where in the last equality of the top line we have used the fact that $\{\rho^a\}$ is the standard basis of $\mathbb{R}^N$. Now, choosing parameters the same as in Fig. (3.7), namely $\hat{m} = 1.0, \hat{g}^2 = 1/2$, as well as $\hat{\mu} = 1.0$ and $N = 8$, we compute the statistical power of the estimator as a function of flow time. One finds that the statistical power is near unity for small flow times and asymptotes to $\Sigma \sim 0.75$ at large flow times. The fact that the statistical power of the estimator remains appreciably different from zero demonstrates its usefulness in this particular case. Moreover, with $\sim 75\%$ of the sample contributing in Fig. (4.1) one needs gather 4/3 as many configurations as when an estimator is not used, but since the cost

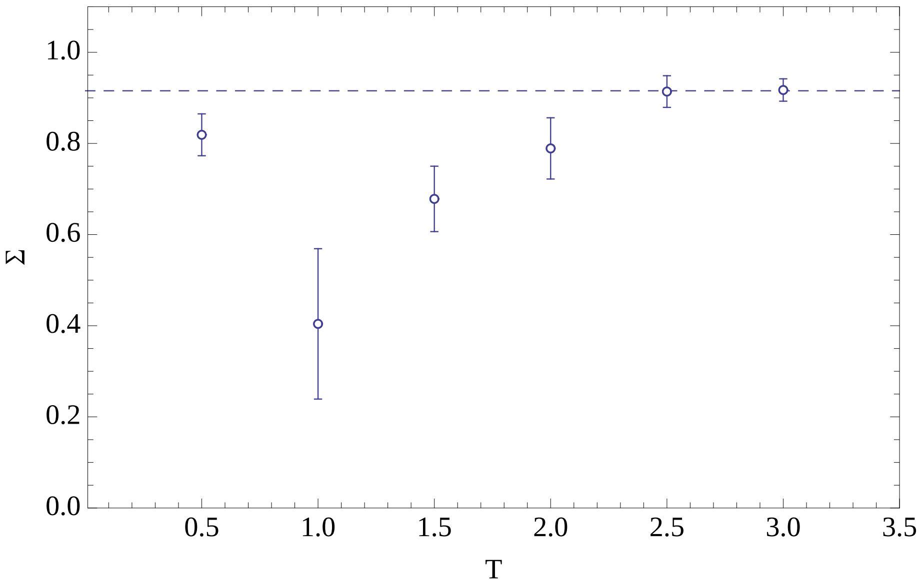

Figure 4.2: Statistical power of reweighting for lattice of size $N = 32$. All parameters used in this plot are the same as the parameters of Fig. (4.1), except the temperature, which is 1/4 of the temperature in Fig. (4.1).

of computing $\ln \det \tilde{J}_1$ is so much smaller than the cost of computing the actual jacobian, simulations are effectively much faster when the estimator is used.

Next let us explore the continuum zero-temperature behavior of the estimator. We take parameters based off those used in Fig. (4.1), which are $N = 8$, $\hat{g}^2 = 1/2$, $\hat{m} = 1.0$. To take the zero temperature limit, all that is necessary is to take $N \to \infty$ while keeping all other parameters fixed. In Fig. (4.2) we have lowered the temperature by a factor of four relative to Fig. (4.1). One can see that there is a dip in the statistical power around $T = 1.0$, where $\Sigma \sim 0.4(2)$, but at all times the statistical power is large enough that the using the estimator remains profitable.

Chapter 5

Application 2: 1+1 Thirring Model

5.1 Continuum theory and Lattice Formulation

In this section we consider the two-flavor (1+1) dimensional Thirring Model at finite density. Adding a dimension makes the problem harder due to the increased number of degrees of freedom, and our purpose is to demonstrate that the holomorphic gradient flow can handle this increase in complexity. To tackle this problem we will use Metropolis proposals and estimators for jacobians discussed in the last chapter to increase the algorithm's speed.

The continuum Lagrangian describing the theory is

$$S = \int dtdx \Big\{ \frac{N_f}{2g^2} A_\mu^2 + \overline{\psi}_\alpha (\not{\partial} + i\not{A} + m + \mu\gamma^0)\psi_\alpha \Big\} \tag{5.1}$$

where $A = (A_0, A_1)$ is a vector field, ψ_α at fixed α is a two component Dirac spinor and α is a flavor index. In this section we elect to analyze the $N_f = 2$ theory, which allows us to use both Wilson and staggered fermions (the staggered formulation can only support an even number of flavors). We will break from the notation of Chapter 3 by denoting fields as A_μ rather than ϕ because our fields now have more than one component.

Similar to the (0+1) case, our lattice action is formulated with compact auxiliary fields and reads in the generic case [21]

$$S = \sum_{x,\nu} \frac{N_F}{g^2}(1 - \cos A_\nu(x)) + \sum_{x,y} \bar{\psi}^\alpha(x) D_{xy}(A) \psi^\alpha(y) \ . \tag{5.2}$$

The fermion matrix using for Wilson fermions is

$$D^W_{xy} = \delta_{xy} + \kappa \sum_{\nu=0,1} \Big[(1-\gamma_\nu) e^{iA_\nu(x)+\mu\delta_{\nu 0}} \delta_{x+\nu,y} - (1+\gamma_\nu) e^{-iA_\nu(y)-\mu\delta_{\nu 0}} \delta_{x,y+\nu} \Big], \tag{5.3}$$

where the Wilson parameter is $\kappa = 1/(2m+4)$ in (1+1) dimensions and where, as before, the fermions are anti-periodic in time $\psi_\alpha(x + N_t\hat{0}) = -\psi_\alpha(x)$ and the auxiliary fields are periodic in time $A_\mu(x + N_t\hat{0}) = A_\mu(x)$. To reach infinite volume physics faster, we choose periodic spatial boundary conditions for both the fermions and the auxiliary fields.

We choose the following gamma matrices

$$\gamma^0 \equiv \sigma_3, \ \gamma^1 \equiv \sigma_1, \ \gamma^5 \equiv \sigma_2 \tag{5.4}$$

The continuum limit of this lattice theory encodes two flavors of Dirac spinors (a Dirac spinor in 1+1 dimensions has two components [107]).

In the staggered fermion action, there is a single Grassmann variable at each site. The staggered fermion matrix reads

$$D^{KS}_{xy} = m\delta_{xy} + \frac{1}{2} \sum_{\nu=0,1} \Big[\eta_\nu(x) e^{iA_\nu(x)+\mu\delta_{\nu 0}} \delta_{x+\nu,y} - \eta^\dagger_\nu(y) e^{-iA_\nu(y)-\mu\delta_{\nu 0}} \delta_{x,y+\nu} \Big] \ , \tag{5.5}$$

where the "staggered phases" $\eta_\nu(x)$ are $\eta_0(x) = 1$ and $\eta_1(x) = (-1)^{x_0}$. Such an action encodes two flavors of two component Dirac fermions in the continuum limit as well. With these definitions, the action after integrating out the fermions reads

$$S = N_f \Big(\sum_{x,\nu} (1 - \cos(A_{\nu x})\ ) - \gamma \ln \det D(A) \Big) \tag{5.6}$$

where $\gamma = 1$ for Wilson fermions and $\gamma = 1/2$ for staggered fermions.

5.2 The Tangent Plane

As in the (0+1) Thirring model, we create our manifold of integration by first shifting the original domain of integration into the complex plane and then flowing it. However, in (1+1) dimensions there are both A_0 and A_1, unlike the (0+1) case where there was only ϕ. One finds that the higher dimensional analog of the constant field critical point is in the imaginary A_0 direction. That is, the action is extremized at constant field point $A_0(x) = i\alpha,\ A_1(x) = 0$ for a particular α which satisfies [21]

$$\frac{\partial}{\partial A_\mu(x)} S(A_0(x), A_1(x))|_{A_0(x)=i\alpha,\ A_1(x)=0} = 0\ . \tag{5.7}$$

Using the action Eq. (5.6), one can write the extremum condition implicitly as the following "gap" equation

$$i \sinh(\alpha) = g^2 \gamma \operatorname{tr} \Big(D^{-1} \frac{\partial D}{\partial A_\mu(x)} \Big) \tag{5.8}$$

One can develop an intuition for this extremum by looking at the action Eq. (5.3). The chemical potential couples only to the timelike links $e^{iA_0(x)+\mu}$. If one makes the substitution $A_0(x) \to A_0(x)+i\mu$, then the phase fluctuations from the determinant of the fermion matrix vanish because the fermion matrix is rendered real. This substitiution renders the gauge term $\sum_x 1 - \cos(A_0(x))$ complex however, and the solution to the gap equation balances the effect of these two terms. One finds that the Takagi vectors attached to this extremum are purely real. This renders

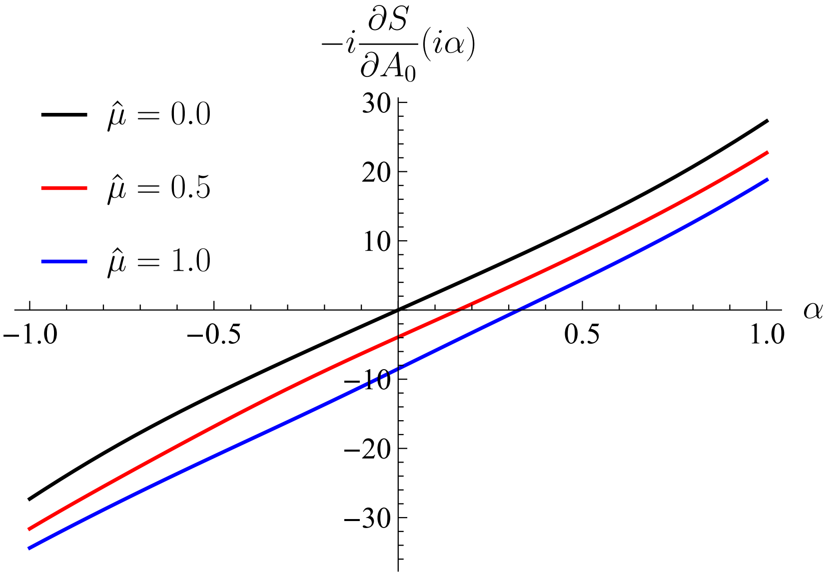

Figure 5.1: Here we plot the derivative of the action for constant field configurations $A_0 = i\alpha$, $A_1 = 0$ at various chemical potentials for Wilson fermions. Here $N_t = N_x = 4$, and $\hat{m} = 0.1$, $\hat{g}^2 = 1.0$. One sees that the extremum of the action coincides with the real plane when $\hat{\mu} = 0$, and it begins to move up in the imaginary direction as $\hat{\mu}$ is increased.

the tangent space of the thimble at this extremum parallel to the real plane, as with the (0+1) case. This once again makes simulations on the tangent plane computationally inexpensive.

We presently solve the condition for the constant field extremum more explicitly. If we consider field configurations constant in spacetime, then the fermion portion of the action can be diagonalized in momentum space. Once the action is diagonalized, taking the determinant is trivial and one can show that for constant fields $(A_0(x), A_1(x)) = (A_0, 0)$, the action (for Wilson fermions) is

$$S/N_f = \frac{N_t N_x}{\hat{g}^2}(1-\cos(A_0)) - \sum_{p_0,p_1} \ln\left[1 + 8\kappa^2 + 4\kappa\cos(A_0 - p_0 - i\mu)\big(1 + 2\kappa\cos(p_1)\big)\right] \tag{5.9}$$

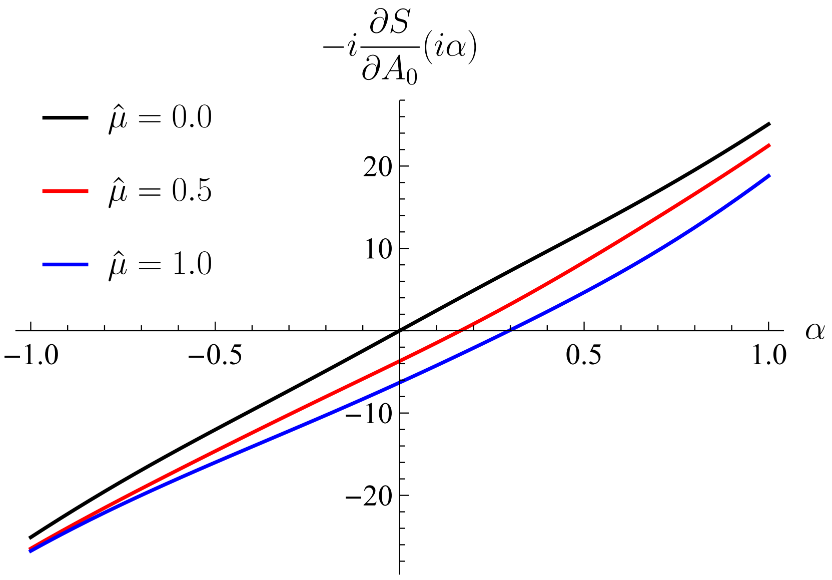

Figure 5.2: Here we plot the derivative of the action for constant field configurations $A_0 = i\alpha,\ A_1 = 0$ at various chemical potentials for staggered fermions. Here $N_t = N_x = 4$, and $\hat{m} = 0.1,\ \hat{g}^2 = 1.0$.

Consequently the gradient of the action is

$$\frac{\partial}{\partial A_0}\Big(S/N_f\Big) = \frac{N_t N_x}{\hat{g}^2}\sin(A_0) + \sum_{p_0,p_1} \frac{4\kappa\sin(A_0 - p_0 - i\mu)\big(1 + 2\kappa\cos(p_1)\big)}{1 + 8\kappa^2 + 4\kappa\cos(A_0 - p_0 - i\mu)\big(1 + 2\kappa\cos(p_1)\big)} \tag{5.10}$$

These sums can be done numerically and the derivative is plotted for $A_0 = i\alpha$ in Fig. (5.1). It is clear from the plot that, at $\hat{\mu} = 0$, the tangent plane is the original domain of integration. As $\hat{\mu}$ is increased, however, the tangent plane rises up along the imaginary axis.

The analysis is much the same with staggered fermions. Assuming once again a constant fields, the fermion term of the staggered action in momentum space is equal to

$$S_f = \sum_{k_0,k_1} \bar{\chi}(k)\Big[m\mathbb{1} + \frac{1}{2}\sigma_3\big(\mathrm{e}^{iA_0+\mu-ik_0} - \mathrm{e}^{-iA_0-\mu+ik_0}\big) + \frac{1}{2}\sigma_1\big(\mathrm{e}^{ik_1} - \mathrm{e}^{-ik_1}\big)\Big]\chi(k) \tag{5.11}$$

where $\chi(k)$ is a two component spinor field with $0 \leq k_0 \leq \pi$ and $0 \leq k_1 \leq 2\pi$ and

$$\chi(k) = \begin{bmatrix} \psi(k_0, k_1) \\ \psi(k_0 + \pi, k_1) \end{bmatrix} \tag{5.12}$$

Integrating out the fermions then yields the effective action

$$S/N_f = \frac{N_t N_x}{\hat{g}^2}(1 - \cos(A_0)) - \frac{1}{2} \sum_{k_0 < \pi, k_1} \ln\Big[m^2 + \sin(A_0 - k_0 - i\mu)^2 + \sin(k_1)^2\Big] \tag{5.13}$$

which leads to the derivative

$$\frac{\partial}{\partial A_0}\Big(S/N_f\Big) = \frac{N_t N_x}{\hat{g}^2}\sin(A_0) - \frac{1}{2} \sum_{k_0 < \pi, k_1} \left[\frac{2\sin(A_0 - k_0 - i\mu)\cos(A_0 - k_0 - i\mu)}{m^2 + \sin(A_0 - k_0 - i\mu)^2 + \sin(k_1)^2}\right] \tag{5.14}$$

Once again this derivative can be evaluated numerically. We plot the derivative along the imaginary A_0 axis in Fig. (5.2). One finds similar behavior for the location of the extremum as in the Wilson case.

5.3 The Flow

In order to flow field configurations it is necessary to know the first derivative of the action. Additionally, to flow either the jacobain or estimators thereof one needs the second derivative of the action. We record these results in this section. Those uninterested in the fine details of the calculation are encouraged to skip ahead to the next section.

For Wilson fermions, the first derivative of the action is

$$\frac{\partial}{\partial A_\alpha(z)}\big(S/N_f\big) = \frac{\sin A_\alpha(z)}{g^2} - \mathrm{tr}\left(D^{-1}\frac{\partial D}{\partial A_\alpha(z)}\right) \tag{5.15}$$

where the derivative of the dirac matrix is

$$\Big(\frac{\partial}{\partial A_\alpha(z)} D\Big)_{xy} = i\kappa\Big[(\mathbb{1}-\gamma_\alpha)\delta_{xz}\delta_{y,x+\alpha}\mathrm{e}^{iA_\alpha(z)+\mu\delta_{\alpha 0}} + (\mathbb{1}+\gamma_\alpha)\delta_{x,z+\alpha}\delta_{y,z}\mathrm{e}^{-iA_\alpha(z)-\mu\delta_{\alpha 0}}\ .\Big] \tag{5.16}$$

The second derivative of the action is

$$\frac{\partial}{\partial A_\beta(z')}\frac{\partial}{\partial A_\alpha(z)}\big(S/N_f\big) = \tag{5.17}$$
$$= \delta_{z,z'}\delta_{\alpha\beta}\frac{\cos A_\alpha(z)}{g^2} + \mathrm{tr}\,\big(D^{-1}\frac{\partial D}{\partial A_\beta(z')}D^{-1}\frac{\partial D}{\partial A_\alpha(z)}\big) - \mathrm{tr}\,\big(D^{-1}\frac{\partial^2 D}{\partial A_\beta(z')\partial A_\alpha(z)}\big)$$

with the second derivative of the dirac matrix

$$\Big(\frac{\partial}{\partial A_\beta(z')}\frac{\partial}{\partial A_\alpha(z)} D\Big)_{xy} = \tag{5.18}$$
$$= (i^2\kappa)\delta_{\alpha,\beta}\delta_{z,z'}\Big[(\mathbb{1}-\gamma_\alpha)\delta_{xz}\delta_{y,x+\alpha}\mathrm{e}^{iA_\alpha(z)+\mu\delta_{\alpha 0}} - (\mathbb{1}+\gamma_\alpha)\delta_{x,z+\alpha}\delta_{y,z}\mathrm{e}^{-iA_\alpha(z)-\mu\delta_{\alpha 0}}\Big]\ .$$

For staggered fermions

$$\frac{\partial}{\partial A_\alpha(z)}\big(S/N_f\big) = \frac{\sin A_\alpha(z)}{g^2} - \frac{1}{2}\,\mathrm{tr}\,\big(D^{-1}\frac{\partial D}{\partial A_\alpha(z)}\big) \tag{5.19}$$

where

$$\Big(\frac{\partial}{\partial A_\alpha(z)} D\Big)_{xy} = \frac{i}{2}\eta_\alpha(z)\Big[\delta_{xz}\delta_{y,x+\alpha}\mathrm{e}^{iA_\alpha(z)+\mu\delta_{\alpha 0}} + \delta_{x,z+\alpha}\delta_{y,z}\mathrm{e}^{-iA_\alpha(z)-\mu\delta_{\alpha 0}}\Big]\ . \tag{5.20}$$

The second derivative of the action with staggered fermions takes the same form as with Wilson fermions with the only difference being the factor of 1/2 in the staggered case. Finally, the second derivative of the Dirac matrix for staggered fermions is:

$$\Big(\frac{\partial}{\partial A_\beta(z')}\frac{\partial}{\partial A_\alpha(z)}D\Big)_{xy} = \tag{5.21}$$
$$= (i^2/2)\delta_{\alpha,\beta}\delta_{z,z'}\eta_\alpha(z)\Big[\delta_{xz}\delta_{y,x+\alpha}\mathrm{e}^{iA_\alpha(z)+\mu\delta_{\alpha 0}} - \delta_{x,z+\alpha}\delta_{y,z}\mathrm{e}^{-iA_\alpha(z)-\mu\delta_{\alpha 0}}\Big]\ .$$

5.4 Results

The results in this section are from [21]. To determine physical parameters, we measure two particle masses: a fermion and a boson. Denoting the lattice spacing as a, we extract am_f and am_b from the long-time behavior of the correlators $\langle \mathcal{O}_f^\dagger(t)\mathcal{O}_f(0)\rangle$ and $\langle \mathcal{O}_b^\dagger(t)\mathcal{O}_b(0)\rangle$ where the fermion interpolating field is $\mathcal{O}_f(t) = \sum_{\vec{x}}\psi_1(\vec{x},t)$ and the boson interpolating field is $\mathcal{O}_b(t) = \sum_{\vec{x}}\bar{\psi}_i(\vec{x},t)\gamma_5(\tau_3)_{ij}\psi_j(\vec{x},t)$. The subscripts label the flavor index of the fields, while the spinor index is left implicit (so that ψ_1 for example is a two component object). Note that for a free theory, $m_b = 2m_f$. We can therefore gauge the strength of the interaction by the ratio m_b/m_f: when this quantity is significantly different from 2 the theory is strongly interacting. For Wilson fermions, we choose $\hat{g} = 1.0$ and $\hat{m} = -0.25$. For such bare parameters, $am_f = 0.30(1)$ and $am_b = 0.44(1)$. Since the boson is bound by about half a fermion mass we conclude that the theory is strongly interacting. For staggered fermions we choose bare parameters so that the physical masses are the same as in the Wilson case.

We first explore the severity of the sign problem in this theory. In Fig. (5.3) we show results obtained on a 10 × 10 lattice with Wilson fermions on various manifolds. In the right half of the figure we plot the average sign, which is a measure of the sign problem; when the average sign is close to 1 the sign problem is gentle, and when the average sign is close to 0 the sign problem is severe. The black points are obtained on the standard domain of integration. One sees that the average sign drops to zero at about a fermion mass; this is a basic manifestation

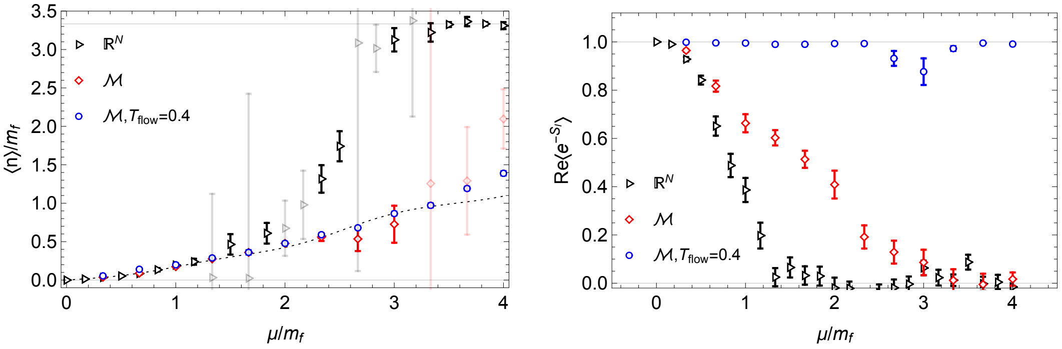

Figure 5.3: Here we plot the density of fermions per flavor (left) and the average sign (right) as a function of the chemical potential in units of the fermion mass. In the left plot, the upper horizontal line is the saturation density and the dotted curve corresponds to the free gas result; we grayed out the points with the errorbars exceeding 0.3 to make the figure easier to read. In the right figure, the black points are calculations on the standard domain of integration where the sign problem is severe, the red points are calculated on a shifted contour, and the blue points are calculated on a shifted and flowed contour. The ever increasing average sign demonstrates that the holomorphic gradient flow tames the sign problem.

of the sign problem. We also note that when $\mu \simeq m_f$, fermions begin to populate the box. It appears that the sign problem quickly worsens when the box begins to be populated. We don't have a solid understanding of why this occurs at this time. The red points come from integrating on the tangent plane to the critical point in the constant field subspace. One sees that the sign problem is tamed for $\mu/m_f \sim 3$, which is a substantial improvement over the tangent plane. A key point to note here is that calculations on the tangent plane cost just as little as calculations on the real plane, and so integrating on a shifted manifold is useful. In order to probe higher chemical potenials one must flow the shifted plane. Flowing the shifted manifold by a flow time $T = 0.4$, one obtains a manifold whose average sign is plotted in blue. It is clear that the sign problem is solved for all values of the chemical potential.

It is helpful to see how the sign problem is reflected in calculations of observables. We present calculations of the density on the real plane (black), tangent plane (red) and flowed manifold (blue) in the left half of Fig. (5.3). The smallness of the average sign for real plane calculations is reflected in the dramatic increase in the error bars on the density after $\mu/m_f \sim 1$. This rapid increase in statistical uncertainty causes one to lose all predictability. One sees similar behavior in the red data: while the average sign is under control and distinguishable from zero the statistical uncertainty on the density is manageable and the Monte Carlos are predictive. However, once the average sign drops to zero at around $\mu/m_f \sim 3$, the error bars blow up and one loses accuracy. Finally, it is clear that the statistical uncertainty is under control for all values of the chemical potential on the flowed manifold, reflecting the fact that on this manifold the sign problem is tamed.

We pause for a moment to remark on the behavior of the error bars in the density plot of Fig. (5.3). It will be noticed that the size of the error bar fluctuates wildly from point to point as the density is increased. As explained in the Introduction, until the average sign is distinguishable from zero, observables are completely unknown and the error bars are correspondingly large.

We further note a useful, albeit loose, rule. By standard error analysis, the statistical uncertainty in any observable $\mathcal{O}$ is proportional to

$$\delta\mathcal{O} \sim \frac{1}{\langle\sigma\rangle^2} \tag{5.22}$$

where σ is the average sign. This is because for a function $f(x,y)$, the relative error due to uncertainties $\delta x, \delta y$ in x, y is given by $\delta f/f = \sqrt{(\delta x/x)^2 + (\delta y/y)^2}$. This means that when the average sign is close to zero, much can be gained from a small improvement in the average sign. For example, if the average sign goes from 0.001 to 0.01 by a suitable improvement of the manifold of integration, then the statistics needed for a specified error bar goes down by a factor of 100.

Having demonstrated that the sign problem is severe at large chemical po-

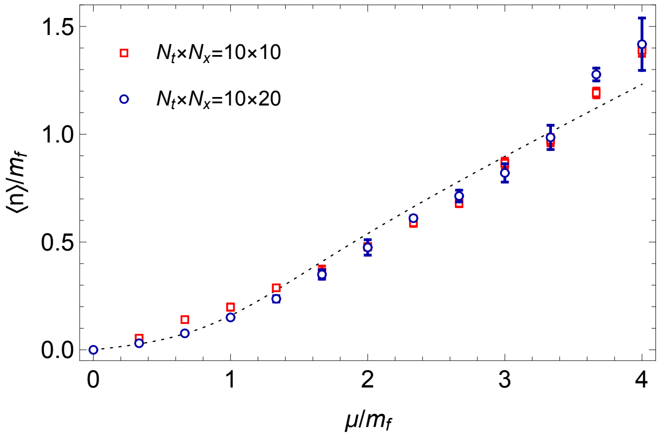

Figure 5.4: Fermion density as a function of chemical potential on two different volumes, 10 × 10 from Fig. (5.3) and 10 × 20 using the same parameters. Dotted line is the free gas result.

tentials and having showed that the holomorphic gradient flow can cure it for a particular set of parameters, let us explore the efficacy of the method for different parameter sets. We will consider the infinite volume, zero temperature and continuum limits in that order.

First we demonstrate that the sign problem is tamed even for lattices close to the thermodynamic limit. As expected from standard arguments [17], we find that the severity of the sign problem increases as the volume grows, however we are able to tame the growing sign problem with the holomorphic gradient flow. In Fig. (5.4) we compute the density for Wilson fermions with the same couplings specified in Fig. (5.3). We calculate on manifolds flowed by $T = 0.4$ for both volumes. It is clear that the sign problem is under control in both calculations.

Next we consider the zero temperature limit and demonstrate that our method can handle temperatures low enough to be deep in the degenerate limit and exhibit the "Silver Blaze" phenomenon [108]. We find that the sign problem worsens as the temperature is decreased, but over the parameters studied the sign problem remained manageable. In Fig. (5.5) we plot the density as a function of the

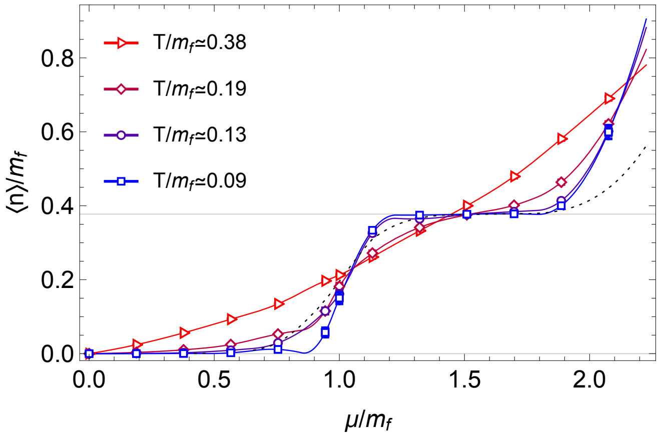

Figure 5.5: $\langle n \rangle$ as a function of μ for several temperatures using staggered fermions. The horizontal line is the density that corresponds to one particle in the box (per flavor). The solid curves are splines interpolations of the data points to guide the eye. The dotted curve represents a fermion free gas in the staggered discretization on a 40×10 with the dimensionless mass set to the value of am_f.

chemical potential for temperatures ranging from "hot" ($T/m_f = 0.38$) in red to "cold" ($T/m_f = 0.09$) in blue. The lattices for these simulations range from 10×10 to 40×10. Two clear plateaus emerge in the cold calculations at $\mu \simeq 1.0$ and $\mu \simeq 2.0$ There are two reasons this is a non-trivial result. First, standard Monte Carlo methods can't reach temperatures where this step structure emerges because the sign problem is too severe. Second, the two step structure indicates that our simulations are not trapped to the main thimble. The holomorphic gradient flow solves the sign problem by gradually creating pockets of high probability and stationary phase, so there is always the worry that a simulation is trapped around a single local minima attached to one thimble. We can be reasonably confident that this is not the case in our simulations however because it has been shown, at least in (0+1) dimensions [109], that an integration around a single thimble leads to a linear increase in the density. The fact that we find a step structure in the density indicates that field space is appropriately sampled. Note that we find staggered fermions to have a less severe sign problem than Wilson fermions,

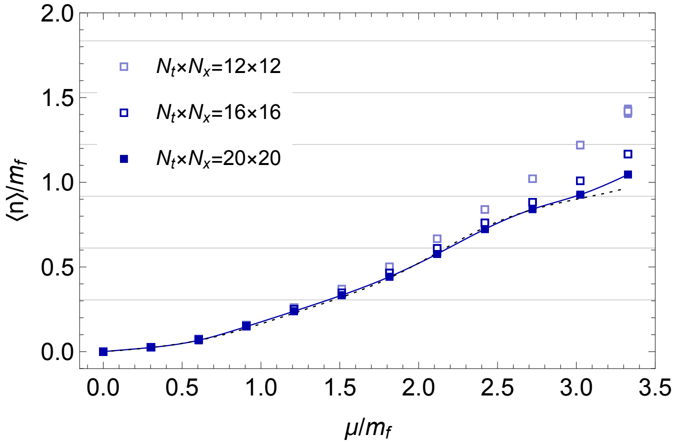

Figure 5.6: Particle density as a function of the chemical potential for different lattice spacings, for fixed volume ($m_f L \approx 3.31$) and temperature ($T/m_f \approx 0.302$). The solid line represents a spline interpolating through finest lattice spacing data points. The dotted line represented the fermion free gas result. Horizontal lines indicate integral number of particles in the box.

so in these cold temperature simulations we use staggered fermions. Furthermore are staggered fermions are 8× cheaper than Wilson fermions. The calculations of Fig. (5.5) were done on the tangent plane.

Finally, we demonstrate that the holomorphic gradient flow can reach a regime “fairly close” to the continuum limit. Again we use staggered fermions and we vary the lattice spacing in the ratios $\frac{20}{20} : \frac{20}{16} : \frac{20}{12}$ while keeping the ratio $m_b/m_f \simeq 1.70$, as well as the temperature and volume fixed. We find that the sign problem varies little with the lattice spacing, rendering the calculations of Fig. (5.6) possible on the tangent plane.

Chapter 6

Sign-Optimized Manifolds

In this chapter we develop an approach for generating manifolds with tame sign problems which is quite different from the holomorphic gradient flow, the so-called "sign-optimized manifold" method [22,48]. Recall that the sign problem makes computing observables difficult because in the reweighting expression

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O} \mathrm{e}^{-i \mathrm{Im}\, S} \rangle_{\mathrm{Re}\, S}}{\langle \mathrm{e}^{-i \mathrm{Im}\, S} \rangle_{\mathrm{Re}\, S}} \tag{6.1}$$

the denominator and the numerator are exponentially small in the spacetime volume. The objective of the sign-optimized method, just like that of the holomorphic gradient flow, is to make the average sign as large as possible. The sign-optimized manifold method is also a deformation of the domain of integration of the path integral (and is therefore correct due to Cauchy's theorem), however the means by which this is achieved is quite different than with the holomorphic gradient flow. Instead of flowing the original domain of integration, one proposes a family of complex manifolds parameterized by a few real parameters, then systematically hunts around in manifold space for the optimal manifold to integrate over. In this chapter we will develop this idea in detail then apply it to the (2+1) Thirring model at finite density in the next.

6.1 The Method

Consider a lattice theory with a sign problem. Call the action S and let the the field variables be denoted as $A_{\mu,x}$. Then on the original domain of integration the average sign $\langle \mathrm{e}^{-i\mathrm{Im}\,S} \rangle_{\mathrm{Re}\,S}$ is exponentially small in the spacetime volume [17]. It can be made larger with the sign-optimized manifold method, which is the following procedure:

1. Propose a family of complex manifolds $\mathcal{M}(\vec{\lambda})$ parameterized by real numbers $\vec{\lambda}$. One typically does this by explicitly specifying the functions $\tilde{A}_{\mu,x} = \tilde{A}_{\mu,x}(A;\vec{\lambda})$.
2. The average sign is locally maximized over the family of manifolds. This is done with a gradient descent algorithm in our work.
3. Observables are calculated on the obtained manifold.

The details of the first step vary from theory to theory because what a "good" manifold looks like depends on the details of the action S. Therefore we will postpone a detailed discussion of this step until we begin the analysis of the (2+1) Thirring Model at finite density. The second step warrants discussion. The average sign, which we denote as $\langle \sigma \rangle \equiv \langle \mathrm{e}^{-i\mathrm{Im}\,S} \rangle_{\mathrm{Re}\,S}$, is a function of $\vec{\lambda}$:

$$\langle \sigma \rangle(\vec{\lambda}) = \frac{\int_{\mathcal{M}(\vec{\lambda})} D\tilde{A}\ \mathrm{e}^{-S(\tilde{A})}}{\int_{\mathcal{M}(\vec{\lambda})} D\tilde{A}\ \mathrm{e}^{-\mathrm{Re}\,S(\tilde{A})}} \tag{6.2}$$

The average sign is very difficult to compute stochastically in general because it is a globally defined quantity, requiring an full integration to calculate, and furthermore it suffers from the sign problem. Luckily, as we will show momentarily, the gradient of the modulus of the average sign is sign-problem free, and is therefore easy to compute. This means, given a family of manifolds, without once computing the average sign of a member, it is possible to find the member with the largest (in modulus) average sign.

To calculate the gradient of the modulus of the average sign it suffices to calculate its logarithm because $\ln(x)$ is monotonically increasing. We therefore need to find the point where $\nabla_\lambda \ln|\langle\sigma\rangle|$ vanishes. To compute this derivative explicitly, it is necessary to include the jacobian associated with the choice of coordinates. We will denote the jacobian matrix $J(A;\lambda) = \frac{\partial}{\partial A}\tilde{A}(A;\vec{\lambda})$. Then

$$\nabla_\lambda \ln|\langle\sigma\rangle| = \frac{\int_{\mathbb{R}^N} DA\ \mathrm{e}^{-\mathrm{Re}\ (S-\ln\ \det J)}\big(\mathrm{Re}\ \nabla_\lambda S(A) - \mathrm{Re}\ \ \mathrm{tr}(J^{-1}\nabla_\lambda J)\big)}{\int_{\mathbb{R}^N} DA\ \mathrm{e}^{-\mathrm{Re}\ (S-\ln\ \det J)}}\ . \tag{6.3}$$

The calculation of Eq. (6.3) is sign problem free: only $\mathrm{e}^{-\mathrm{Re}\ (S-\ln\ \det J)}$ appears in the Bolzmann weight. Using Eq. (6.3) we locally minimize the modulus of the average sign using a stochastic gradient ascent alogorithm [110]. In practice this is done as follows. First one chooses an initial set of parameters $\vec{\lambda}_1$. Then the gradient, being sign-problem free, is estimated by computing Eq. (6.3) with a short Monte Carlo. Based on this estimate of the gradient, a small step is taken, giving a new set of parameters $\vec{\lambda}_2$. This process is repeated many times, causing the modulus of the average sign to gradually increase. Once the average sign reaches a plateau, say at the parameters $\vec{\lambda}_N$, the process is terminated and the manifold $\mathcal{M}(\vec{\lambda}_N)$ is used for calculating observables.

Chapter 7

Application 3: 2+1 Thirring Model

7.1 Continuum Theory and Lattice Action

We now apply the sign-optimized manifold methd to the (2+1) Thirring model at finite density, whose continuum action is

$$S = \int dt d^2 x \Big\{ \frac{N_f}{2g^2} A_\mu^2 + \overline{\psi}_\alpha (\partial\!\!\!/ + i A\!\!\!/ + m + \mu\gamma^0)\psi_\alpha \Big\} \ . \tag{7.1}$$

As with the (1+1) dimensional model, we choose $N_f = 2$. Little is known about the finite density properties of this model. To my knowledge, beside the calculation presented in this chapter, the only other lattice study of the (2+1) dimensional Thirring Model at finite density is presented in [111], where Complex Langevin is used. While our goal here is to try to solve as difficult sign problems as possible, for the purpose of developing robust techniques, the $(2 + 1)$ Thirring model at finite density is also phenomenologically interesting, appearing in effetive models of high-T_c superconductors and graphene (see references in [112]).

Our lattice action is a direct generalization of the staggered action from (1+1) dimensions and reads:

$$\begin{aligned}S = &\sum_{x\mu} \frac{N_F}{g^2}(1 - \cos A_\mu(x)) + m \sum_x \bar{\psi}_x \psi_x \\ &+ \frac{1}{2} \sum_{x\mu} \eta_\mu(x) \left[\bar{\psi}_x \mathrm{e}^{iA_\mu(x) + \mu\delta_{\mu 0}} \psi_{x+\mu} - \bar{\psi}_{x+\mu} \mathrm{e}^{-iA_\mu(x) - \mu\delta_{\mu 0}} \psi_x \right]\end{aligned} \tag{7.2}$$

where x ranges over a 3 dimensional hypercbic lattice, $\nu = 0, 1, 2$, $\eta_\mu(x) = (-1)^{x_0 + \ldots + x_{\mu-1}}$ and ψ_x is a single component grassman variable at lattice site x. The family of manifolds we propose is

$$\begin{aligned}\tilde{A}_{0,x} &= A_{0,x} + i(\lambda_1 + \lambda_2 \cos A_{0,x} + \lambda_3 \cos 2A_{0,x}), \\ \tilde{A}_{1,x} &= A_{1,x}, \\ \tilde{A}_{2,x} &= A_{2,x} \ .\end{aligned} \tag{7.3}$$

For compact notation let us call $f(A_{0,x}) = \lambda_1 + \lambda_2 \cos A_{0,x} + \lambda_3 \cos 2A_{0,x}$ so that $\tilde{A}_{0,x} = A_{0,x} + if(A_{0,x})$. Notice that we are choosing to complexify only the timelike links, and furthermore the complex variable $\tilde{A}_{0,x}$ only depends on $A_{0,x}$, and not on $A_{0,y}$ with $y \neq x$. This means that the proposed family of manifolds is separable and that the jacobian is diagonal

$$J_{xy,\mu\nu} = \delta_{xy}\delta_{\mu\nu}\big(if'(A_{0,x})\delta_{\mu 0} + \delta_{\mu 1} + \delta_{\mu 2}\big) \ . \tag{7.4}$$

7.2 Proving Sufficient Conditions

For Eq. (7.3) to be a legitimate deformation, our proposed family of manifolds must satisfy the stipulations of Cauchy's theorem. This means that it has to be possible to smoothly deform the original domain of integration $(S_1)^N$ to the final manifold without crossing any poles and without changing the behavior of the integrand at the boundary.

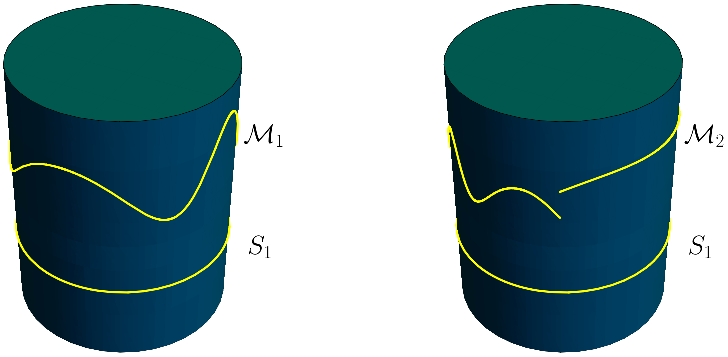

Figure 7.1: Here we repropduce Fig. (7.1) to guide the eye. Deformed manifolds must be continuous for Cauchy's Theorem to hold.

Consider

$$\tilde{A}(A_{0,x})_t = A_{0,x} + tif(A_{0,x}) \tag{7.5}$$

Then at $t = 0$ we have the original manifold and at $t = 1$ we have the final manifold. Furthermore the deformation is smooth because it's indexed by a linear function. Also, since the original domain of integration is $(S_1)^N$, it is necessary that the proposed family of manifolds are 2π periodic. This periodicity ensures that the manifold is not ripped at the edge as t increases, which must hold for the deformation of the manifold to be smooth.

Clearly the deformation Eq. (7.3) is 2π periodic. Indeed, one may interpret the function $f(A_{0,x}) = \lambda_1 + \lambda_2 \cos A_{0,x} + \lambda_3 \cos 2A_{0,x}$ as the most general expansion of a 2π periodic function up to third order in a Fourier series. Finally, no poles are crossed during the deformation process because the integrand $\exp(-\sum_{\mu x}(1 - \cos A_{\mu,x}))\det D(A)$ has no poles in the complexified field space.

7.3 Motivation for Manifolds

The form of the manifold Eq. (7.3) is motivated by more than just satisfying Cauchy's theorem. One immediately sees that the average sign for a sign-optimized manifold *must* be larger than the average sign on the standard domain of integration. This is because the original domain of integration is included in the proposed family of manifolds. Now from the (1+1) dimensional Thirring model we know that a constant vertical shift in the imaginary direction of the A_0 variables alleviates a fair amount of the sign problem. This is the origin of the λ_1 term. What is interesting is that the λ_1 value which is obtained in the gradient ascent is not equal to the vertical shift required to place the integration domain parallel to the tangent space of the constant field extremum [22]. There is a small difference between λ_1 and the position of the tangent plane. We interpret this result as an entropy effect; it is not quite optimal to simply shift the domain of integration to the extremum of the action, fluctuations of the field about this minima (entropy) can be incorporated by an adjusted shift in the imaginary direction.

Next, note that we propose a separable family of manifolds. By separable we mean that $\mathcal{M}(\lambda)$ can be written as a direct product of smaller dimensional manifolds (in the same way that $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$). We do this for several reasons. First, the jacobian associated with the manifold is diagonal and therefore easy to compute and the log determinant of Eq. (7.4) can be written exactly

$$\ln \det J(A; \vec{\lambda}) = \sum_x \ln(1 + i f'(A_{0,x})) \tag{7.6}$$

This is in stark contrast with the holomorphic gradient flow where the jacobian cannot be written analytically and is very expensive to compute. We will see that while the jacobian associated with the family of manifolds are inexpensive to compute, these manifolds are generally less efficient at taming the sign problem than those generated with the holomorphic gradient flow. There are two competing factors, the cost of the jacobian and the improvement of the average sign. We find that, at least in the case of the (2+1) Thirring model (which has a severe sign problem) the balance between the expense of the jacobian and effectiveness of the manifold is in favor of the sign-optimized manifolds.

In addition to computational convenience, there is a theoretical motivation for proposing separable manifolds. In the limit $\mu \to \infty$ where the lattice is saturated with particles, the path integral factorizes [111]:

$$Z = \int DA \; \mathrm{e}^{-S} \xrightarrow[\mu\to\infty]{} \Big(\int d\phi \; \mathrm{e}^{-S_1} \Big)^N \tag{7.7}$$

where S_1 is a function which we'll write down in a moment and N is the number of lattice sites. To see that the path integral does factorize, consider the action

$$\begin{aligned} S = & \sum_{x\mu} \frac{N_F}{g^2}(1 - \cos A_\mu(x)) + m \sum_x \bar{\psi}_x \psi_x \\ & + \frac{1}{2} \sum_{x\mu} \eta_\mu(x) \big[\bar{\psi}_x \mathrm{e}^{iA_\mu(x) + \mu\delta_{\mu 0}} \psi_{x+\mu} - \bar{\psi}_{x+\mu} \mathrm{e}^{-iA_\mu(x) - \mu\delta_{\mu 0}} \psi_x \big] \; . \end{aligned} \tag{7.8}$$

In the limit that μ becomes large, the action can be approximated to exponential precision by

$$S = \sum_{x\mu} \frac{N_F}{g^2}(1 - \cos A_\mu(x)) + \frac{1}{2} \sum_x \bar{\psi}_x \mathrm{e}^{iA_0(x) + \mu} \psi_{x+0} \tag{7.9}$$

in which case the determinant of the dirac matrix can be written exactly

$$\det D(A) = \prod_{\vec{x}} \exp \Big(\sum_t (iA_{0,x} + \mu) \Big) \tag{7.10}$$

and so after integrating out the fermions the action reads

$$S \simeq \sum_x \frac{N_f}{g^2}(1 - \cos(A_{0,x})) - (iA_{0,x} + \mu) + \frac{N_f}{g^2}(1 - \cos(A_{1,x})) + \frac{N_f}{g^2}(1 - \cos(A_{2,x})) \equiv \sum_x S_1(A_{\mu,x}) \tag{7.11}$$

proving the claim that the action is separable.

Another beneficial byproduct of the fact that we can write ln det$J$ explicitly, and a primary motivation for proposing separable manifolds is that we are able to use a hybrid Monte Carlo (HMC) algorithm when we simulate this model. In the HMC algorithm, it is necessary to frequently calculate the derivative

$$\frac{\partial}{\partial A}\big(S(\tilde{(}A)) - \ln\,\det J(A;\vec{\lambda})\big)\ . \tag{7.12}$$

Computing this derivative involves knowing $\frac{\partial}{\partial A}J(A;\vec{\lambda})$, which is a relatively involved procedure on flowed manifolds [92]. On the other hand, the derivative of the jacobian Eq. (7.4) can be written explicitly

$$\frac{\partial}{\partial A_{\rho z}} J_{xy,\mu\nu} = \delta_{\rho 0}\delta_{zx}\left[\delta_{xy}\delta_{\mu\nu}(if''(A_{0,x})\delta_{\mu 0})\right] \tag{7.13}$$

The ability to switch from Metropolis to HMC sampling yields a large speed-up.

We conclude this section by recording the explicit expression of $\nabla_\lambda \mathrm{Re}\,\big(S - \ln\,\det J\big)$, which is needed to carry out the manifold optimization:

$$\begin{aligned}
&\nabla_\lambda \mathrm{Re}\,\big(S(\tilde{A}(A)) - \ln\,\det J(A)\big) \qquad\qquad (7.14)\\
&= \mathrm{Re}\left[\frac{\partial S}{\partial A}(\tilde{A}(A))\cdot\nabla_\lambda A - \nabla_\lambda \sum_x \ln(1+if'(A_{0x}))\right]\\
&= \mathrm{Re}\begin{bmatrix} i\sum_x \frac{\partial S}{\partial A_{0x}}(\tilde{A}(A)) \\ i\sum_x \cos(A_{0x})\frac{\partial S}{\partial A_{0x}}(\tilde{A}(A)) \\ i\sum_x \cos(2A_{0x})\frac{\partial S}{\partial A_{0x}}(\tilde{A}(A)) \end{bmatrix} - \mathrm{Re}\begin{bmatrix} \sum_x \frac{i}{1+if'(A_{0x})} \\ \sum_x \frac{i\ \cos(A_{0x})}{1+if'(A_{0x})} \\ \sum_x \frac{i\ \cos(2A_{0x})}{1+if'(A_{0x})} \end{bmatrix}\ .
\end{aligned}$$

7.4 Hybrid Monte Carlo on Sign-Optimized Manifolds

In this section we detail our Hybrid Monte Carlo algorithm for integrating on sign optimized manifolds. To begin, recall that Cauchy's theorem gives the following set of equalities:

$$\begin{aligned} Z &= \int_{(S_1)^N} DA \ \mathrm{e}^{-S(A)} \\ &= \int_{\mathcal{M}} D\tilde{A} \ \mathrm{e}^{-S(\tilde{A})} \text{ (Cauchy's Theorem)} \\ &= \int_{(S_1)^N} DA \ \det J(A) \mathrm{e}^{-S(\tilde{A}(A))} \text{ (choose coordinates),} \end{aligned} \tag{7.15}$$

where the manifold is defined in coordinates in Eq. (7.3), and the parameterization manifold is $(S_1)^N$. We wish to integrate over this sign-optimized manifold using the reweighting technique. This means we want to sample fields according to

$$p(A) = \frac{\mathrm{e}^{-\mathrm{Re}\ S(A)} |\det J(A)|}{\int DA \ \mathrm{e}^{-\mathrm{Re}\ S(A)} |\det J(A)|} \equiv \frac{\mathrm{e}^{-\mathrm{Re}\ S_{eff}(A)}}{\int DA \ \mathrm{e}^{-\mathrm{Re}\ S_{eff}}} \ , \tag{7.16}$$

where we have defined the effective action $S_{eff}(A) = S(\tilde{A}(A)) - \ln \det J(A)$, and then compute observables from

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O} \mathrm{e}^{-i \mathrm{Im}\ S_{eff}} \rangle_{\mathrm{Re}\ S_{eff}}}{\langle \mathrm{e}^{-i \mathrm{Im}\ S_{eff}} \rangle_{\mathrm{Re}\ S_{eff}}} \ . \tag{7.17}$$

An efficient way to sample configuration space according to Eq. (7.16) is the "Hybrid Monte Carlo" (HMC) method [113] a standard method in lattice QCD. In this approach, one defines a hamiltonian by adding to the action (or in this case Re S) a kinetic term with new auxiliary momentum variables, then generates proposals by alternating between momentum and field updates which are generated by discretized Hamiltonian dynamics. The new configuration generated in this way is distant in field space yet has a high probability of acceptance. This makes for a markov chain which decorrelates quickly, and therefore results in an efficient sampling of field space.

To define an HMC algorithm on our sign-optimized manifold we first define the Hamiltonian

$$H(\pi,A)=\frac{1}{2}\sum_{x\mu}\pi_{x\mu}(J(A)^{-1\dagger}J(A)^{-1})_{xx',\mu\mu'}\pi_{x'\mu'}+\text{Re }S(\tilde{A}(A))\ , \tag{7.18}$$

where J is the jacobian Eq. (7.4). If (π, A) space is then sampled according to the Bolzmann distribution

$$p(\pi,A)=\frac{\text{e}^{-H(\pi,A)}}{\int D\pi DA\ \text{e}^{-H(\pi,A)}} \tag{7.19}$$

then the distribution of interest $p(A)=\frac{\text{e}^{-\text{Re }S_{eff}(A)}}{\int DA\ \text{e}^{-\text{Re }S_{eff}}}$ can be obtained by a marginalization over π. To see this, note first that

$$\int D\pi\ \exp(-\frac{1}{2}\sum_{x\mu}\pi_{x\mu}(J(A)^{-1\dagger}J(A)^{-1})_{xx',\mu\mu'}\pi_{x'\mu'})=\sqrt{\pi}^{N}|\det J(A)| \tag{7.20}$$

Therefore

$$\int D\pi\ p(\pi,A)=\frac{\int D\pi\ \text{e}^{-H(\pi,A)}}{\int D\pi DA\ \text{e}^{-H(\pi,A)}}=\frac{\text{e}^{-\text{Re }S(A)}|\det J(A)|}{\int DA\ \text{e}^{-\text{Re }S(A)}|\det J(A)|}=p(A)\ . \tag{7.21}$$

Our task, then, is to sample (π, A) space according to the distribution Eq. (7.19) then throw out the π configurations collected. Doing so will marginalize over the momentum distribution, yielding the desired distribution Eq. (7.16) over fields A.

To carry this program out we must first sample according the Bolzmann distribution Eq. (7.19). Let us us first write out Eq. (7.22) more explicitly:

$$\begin{aligned}H(\pi,A)&=\sum_{x}\big[\frac{\pi_{0x}^2}{1+f'(A_{0x})^2}+\pi_{1x}^2+\pi_{2x}^2\big]+\text{Re }S(\tilde{A}(A))\\&\equiv K(\pi,A)+U(A)\end{aligned} \tag{7.22}$$

Our HMC algorithm is the following:

1. Begin with configuration (π, A)

2. Draw π' from the distribution $p(\pi') = \frac{\mathrm{e}^{-\frac{1}{2}\pi'^T (J^\dagger J)^{-1}\pi'}}{\int D\pi' \ \mathrm{e}^{-\frac{1}{2}\pi'^T (J^\dagger J)^{-1}\pi'}}$ and replace the configuration by (π', A).

3. Generate a new configuration $(\pi', A) \to (\pi'', A'')$ with any time reversible symplectic integrator of the Hamiltonian equations of motion.

4. Accept/reject final configuration (π_f, A_f) with probability $\min\{1, \mathrm{e}^{-(H(\pi_f, A_f) - H(\pi, A))}\}$.

We will relegate to the appendices a proof that the procedure above generates a Markov chain in (π, A) space which converges to the distribution Eq. (7.19). The details of how we implement steps 2 & 3 of the algorithm are also presented in the appendices.

7.5 Results

We now present calculations of the thermodynamics of the (2+1) Thirring Model using the sign optimized manifold method. Here we take $N_f = 2$ and we set $m = 0.01$ and $g = 1.08$ in the staggered action Eq. (7.2). With these bare parameters, we fix the physical scales by measuring the masses of a fermion and a boson, defined to be the lowest energy excitation coupled to the interpolating operators $\langle \mathcal{O}_f(t)\mathcal{O}_f(0)^\dagger \rangle$ and $\langle \mathcal{O}_b(t)\mathcal{O}_b(0)^\dagger \rangle$, where $\mathcal{O}_f(t) = \sum_{\vec{x}} \psi(\vec{x}, t)$ and $\mathcal{O}_b(t) = \sum_{\vec{x}} (-1)^{x_0+x_1+x_2} \bar{\psi}(\vec{x}, t)\psi(\vec{x}, t)$. The operator $\mathcal{O}_b$ can be seen to be a pseudoscalar, equivalent to $\mathcal{O} = \bar{\psi}\gamma^5\psi$ in the continuum, by noting that $\Gamma^5_{xy} = (-1)^{x_0+x_1+x_2}\delta_{xy}$ is the staggered implementation of the γ^5 matrix. Computing masses on a spatial volume of $L^2 = 10^2$ we find $m_f = 0.46(1)$ and $m_b = 0.21(1)$. This indicates that the system is strongly coupled since the boson is so bound that its mass is smaller than the fermion.

First we demonstrate that the optimization procedure converges to an optimal manifold. In Fig. (7.2) we plot the evolution of parameters $(\lambda_1, \lambda_2, \lambda_3)$ which de-

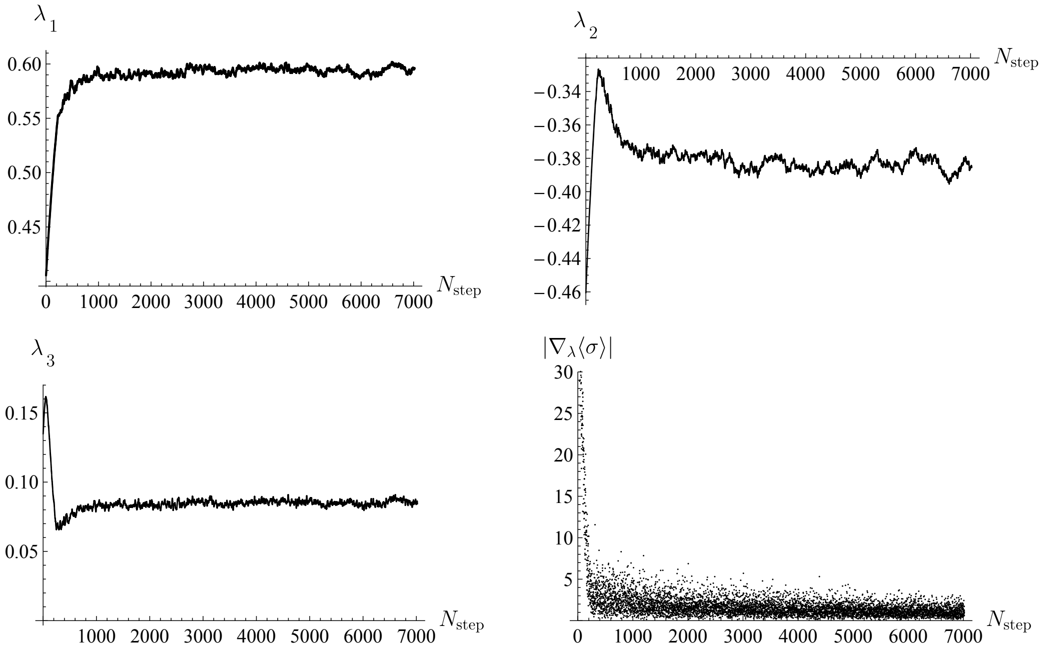

Figure 7.2: Here we plot the evolution of manifold parameters $(\lambda_1, \lambda_2, \lambda_3)$ as a function of the number of steps N_{step} of the optimization procedure. Here $m = 0.01$, $g = 1.08$, $\mu = 0.80$ and the lattice size is 4^3. It can be seen that all three parameters converge to an approximately fixed value, with the stochastic fluctuations in the asymptotic value small compared to the parameters themselves.

fine the sign-optimized manifold for $\mu = 0.80$ on a 4^3 lattice as a function of the optimization step. We randomly initialize parameters and find that, independent of the initial set of parameters the simulation converges to a locally optimal manifold over the course of a simulation. The evolution of parameters in Fig. (7.2) is typical for the range of thermodynamic conditions studied.

We now quantify the improvement of the average sign on our sign-optimized manifolds. In Fig. (7.3) we plot $\log |\langle\sigma\rangle|$ as a function of the chemical potential for several values of the inverse temperature β. In black is the average sign in a $10 \times 6 \times 6$ lattice, computed on the original domain of integration $(S_1)^N$. For small values of the chemical potential, where the average sign is large enough to be easily computed, we find and exponential decay with the chemical potential. This rapid decay of the average sign is why we cannot use the original domain of integration to

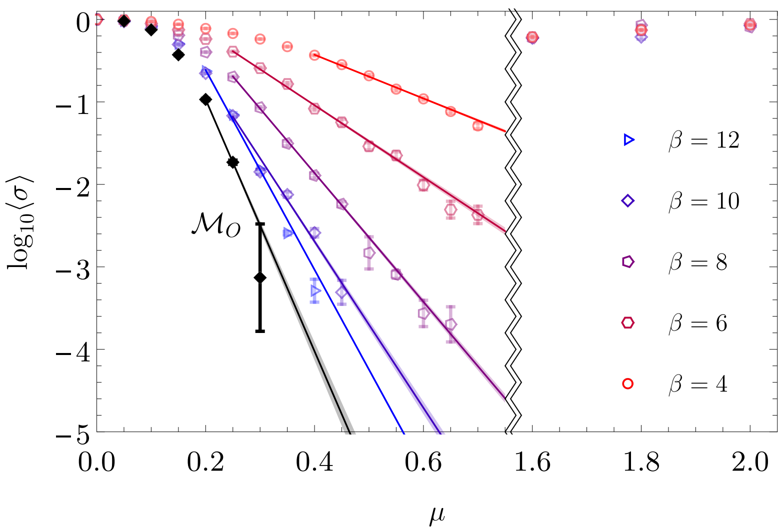

Figure 7.3: Plot of the log-magnitute of the average sign as a function of the chemical potential for inverse temperatures $\beta = 2, 4, ..., 12$ on $\beta \times 6 \times 6$ lattices. The black points are a calculation of the average sign on the original domain of integration $(S_1)^N$ on a $10 \times 6 \times 6$ lattice. We compute for chemical potentials between $0.0 \leq \mu \leq 0.7$, then plot for substantially larger chemical potentials $\mu = 1.6, 1.8.2.0$. The revival of the average sign at large chemical potential is a result of the fact that the path integral factorizes for high μ, which increases the effectiveness of our separable sign-optimized manifold.

study this model. The sign is larger however on the sign-optimized manifolds. For any temperature there is still an exponential decay of the average sign, however the rate of this decay is smaller on sign optimized manifolds. This means that sign-optimized manifolds are exponentially better than $(S_1)^N$ for integration in terms of the statistics needed to compute observables with a specified error, however for large μ computing observables still requires high statistics. For example, the simulations in Fig. (7.3), required between 10^2 and 10^8 configurations. Due to the speed with which the Jacobian can be calculated and the HMC on the manifold, these calculations were done with computational power equivalent to ∼ 20 laptops run over a few weeks.

With the sign problem under control we can now calculate observables. Let

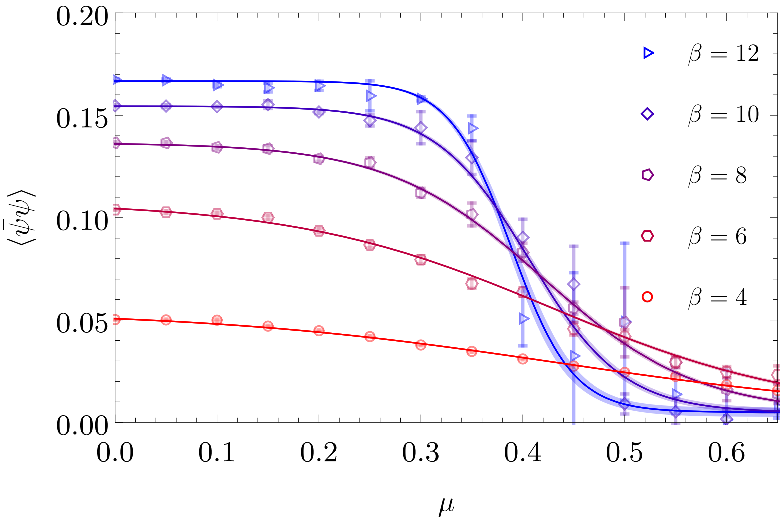

Figure 7.4: Here we plot the chiral condensate $\langle\bar{\psi}\psi\rangle$ as a function of the chemical potential for the same ensembles used in Fig. (7.3). We see the typical behavior of a chiral transition: as the temperature is decreased the condensate develops, and as the chemical potential is increased the condensate decays. The steepness of this transition is sharpened as the temperature is decreased. The solid curves are interpolations of the form $\langle\bar{\psi}\psi\rangle(\mu) = c_0 + c_1\tanh[c_2(\mu - c_3)]$.

us explore the chiral transition of this model which has until now been out of reach. Since the fermion in the theory has a non-zero mass ($am_f = 0.46(1)$) we expect the chiral condensate $\langle\bar{\psi}\psi\rangle \neq 0$ at zero temperature and zero chemical potential. Looking at Fig. (7.4) we find this to be the case. Here we use the same ensembles as in Fig. (7.3); we are computing on $\beta \times 6 \times 6$ lattices. As the temperature decreases at zero chemical potential the chiral condensate develops. Having developed a condensate, as the chemical potential is increased there is an abrupt change in the chiral condensate from a region $\langle\bar{\psi}\psi\rangle \neq 0$ at low μ to a region where $\langle\bar{\psi}\psi\rangle \simeq 0$ at high μ which indicates the presence of a chiral transition in the (2+1) Thirring model.

The calculations of Fig. (7.4) can be converted into a phase diagram, which is plotted in Fig. (7.5). The plot is a heat map of a smooth interpolation of the data in Fig. (7.4) where red denotes high chiral condensate and blue/purple denotes

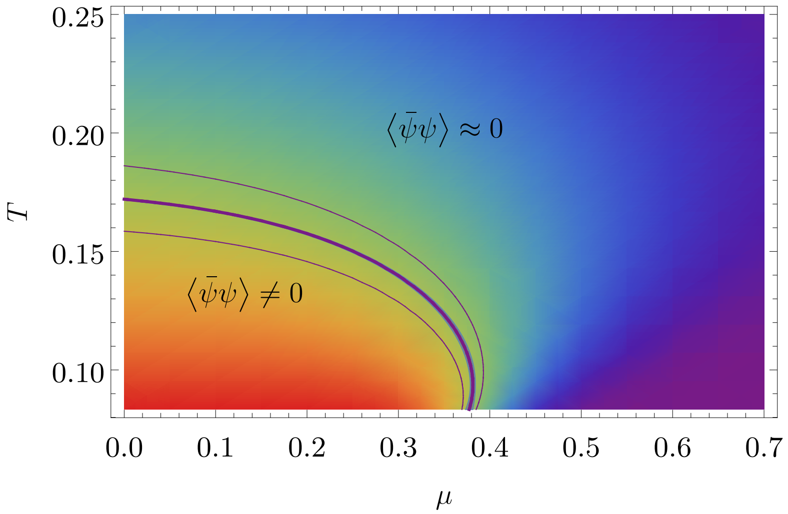

Figure 7.5: Here we plot a heat map of a smooth interpolation of the data in Fig. (7.4). Red denotes high chiral condensate while blue/purple denote little chiral condensate. The purple curve encircling the lower left quadrant is an interpolation between points where $\langle\bar{\psi}\psi\rangle$ is 0.5 of its maximum value (at $\mu = 0$) at a given temperature. The peripheral lines around the thick purple lines interpolate between points where $\langle\bar{\psi}\psi\rangle = 0.45, 0.55$ of its maximum value. While we have not performed continuum limit extrapolation, it is clear that some features of the expected QCD transition are captured in this model.

little chiral condensate. We have not done a systematic study of the continuum limit, however it is clear that some expected features of the QCD phase transitions appear in this model.

Since we have an explicit fermion mass, a second order chiral transition is not expected in this model because the theory is not chirally symmetric. However, one may wonder if at non-zero am_f this transition is first order. An indication of a first order transition would be a sharpening of the jump in the chiral condensate as the volume grows. To that end, we compute on a pair of lattices $10 \times 10 \times 10$ and $10 \times 6 \times 6$ at fixed temperature. The jump in the chiral condensate can clearly be seen to sharpen as the volume increases in Fig. (7.6). The sign problem is extremely severe on the $10 \times 10 \times 10$ lattice and it remains for future work to devise strategies to tackle larger lattices and colder temperatures.

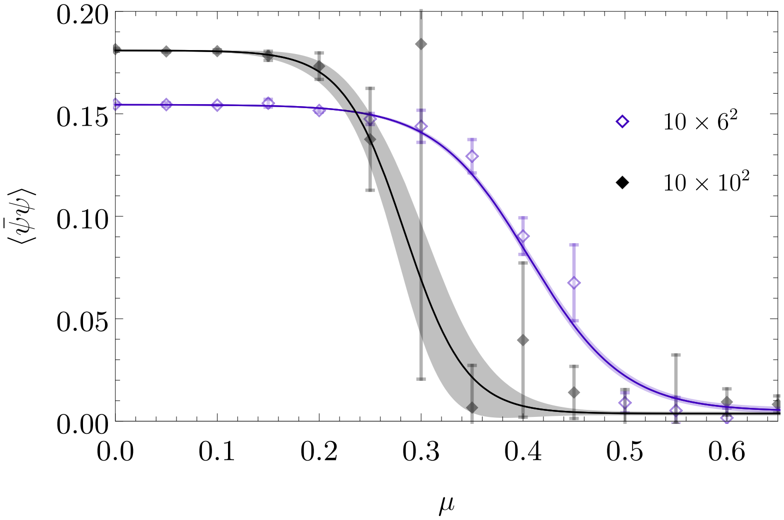

Figure 7.6: Plot of the chiral condensate $\langle\bar{\psi}\psi\rangle$ on $10 \times 10 \times 10$ and $10 \times 6 \times 6$ lattices. The chiral transition is seen to sharpen as the volume is increased, hinting at the existence of a first order transition in this model.

Chapter 8

Appendix

In this appendix we record details of our Hybrid Monte Carlo algorithm used to integrate on sign-optimized manifolds. We first demonstrate that our proposal method satisfies detailed balance. We then demonstrate that our integrator used for Hamiltonian dynamics is symplectic. Finally, we record specific details of the integrator in the case of the (2+1) Thirring model. The details are non-trivial because the manifold of integration is curved.

8.1 Detailed Balance

It is a standard result in Markov Chain Monte Carlo that, to prove that the Markov chain converges to the canonical distribution with respect to the Hamiltonian Eq. (7.22), it is sufficient (but not necessary) to demonstrate that the transition probability that defines the Markov Chain satisfies detailed balance [114]. That is, it suffices to demonstrate

$$\mathrm{Prob}(\pi, A) T((\pi, A) \to (\pi_f, A_f)) = \mathrm{Prob}(\pi_f, A_f) T((\pi_f, A_f) \to (\pi, A)) \tag{8.1}$$

where $\mathrm{Prob}(\pi, A) = d\pi dA \; \frac{\mathrm{e}^{-H(\pi,A)}}{Z}$ is the probability that the fields lie in a phase space volume $d\pi dA$ centered around (π, A) and where $T((\pi, A) \to (\pi_f, A_f))$ is the probability of transitioning from (π, A) to (π_f, A_f). Our transition probability is the composition of two transitions:

$$T = T_2 \circ T_1 \tag{8.2}$$

where

$$T_1((\pi, A) \to (\pi_f, A)) = \frac{\mathrm{e}^{-\frac{1}{2}\pi_f{}^T (J^\dagger J)^{-1}\pi_f}}{\int D\pi'_f \ \mathrm{e}^{-\frac{1}{2}\pi_f{}^T (J^\dagger J)^{-1}\pi_f}} \tag{8.3}$$

is the probability of transitioning from (π, A) to (π_f, A) due to T_1. Notice that this step only changes the momenta. The transition probability T_2 is defined as

$$T_2((\pi, A) \to (\pi_f, A_f)) = \delta((\pi_f, A_f) - \Phi(\pi, A)) \times \min\{1, \frac{\mathrm{e}^{-H(\pi_f, A_f)}}{\mathrm{e}^{-H(\pi, A)}}\} \tag{8.4}$$

where δ is a $2N$ dimensional delta function on phase space and where Φ is the symplectic reversible map from phase space to phase space. Note that by reversibility we mean Φ satisfies

$$\Phi(\Phi(\pi, A)) = (\pi, A) \tag{8.5}$$

Symplecticness means that, if we denote the jacobain matrix of Φ as $D\Phi$, then $D\Phi$ satisfies the relation

$$D\Phi^T j D\Phi = j \tag{8.6}$$

where j is the canonical 2-form on phase space $j = \begin{bmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix}$ [115]. One of the properties of a symplectic transformation is that it leaves volumes of phase space unchanged. That is if $d\pi dA$ is a patch of phase space and of $d\pi' dA'$ is the image of this patch under Φ, then

$$d\pi' dA' = d\pi dA \tag{8.7}$$

That this holds for a symplectic transformation is easily seen from Eq. (8.6): taking the determinant of both sides one finds

$$\det D\Phi^2 = \det j^2 = 1 \tag{8.8}$$

Now, if each transition probability satisfies detailed balance, then so does their composition. Note that

$$p(\pi, A)T_1((\pi, A) \to (\pi_f, A)) = d\pi dA \frac{\mathrm{e}^{-H(\pi,A)}}{Z} \frac{\mathrm{e}^{-\frac{1}{2}\pi_f{}^T(J(A)^\dagger J(A))^{-1}\pi_f}}{\int D\pi'_f \ \mathrm{e}^{-\frac{1}{2}\pi_f{}^T(J(A)^\dagger J(A))^{-1}\pi_f}} \tag{8.9}$$

and

$$p(\pi_f, A)T_1((\pi_f, A) \to (\pi, A)) = d\pi dA \frac{\mathrm{e}^{-H(\pi_f,A)}}{Z} \frac{\mathrm{e}^{-\frac{1}{2}\pi^T(J(A)^\dagger J(A))^{-1}\pi}}{\int D\pi' \ \mathrm{e}^{-\frac{1}{2}\pi^T(J(A)^\dagger J(A))^{-1}\pi}} \tag{8.10}$$

so T_1 satisfies detailed balance. To see that T_2 satisfies detailed balance, we once again compute both sides of the detailed balance condition:

$$p(\pi, A)T_2((\pi, A) \to (\pi_f, A_f)) = d\pi dA \frac{\mathrm{e}^{-H(\pi,A)}}{Z} \delta((\pi_f, A_f) - \Phi(\pi, A)) \min\{1, \frac{\mathrm{e}^{-H(\pi_f,A_f)}}{\mathrm{e}^{-H(\pi,A)}}\} \tag{8.11}$$

and

$$p(\pi_f, A_f)T_2((\pi_f, A_f) \to (\pi, A)) = d\pi_f dA_f \frac{\mathrm{e}^{-H(\pi_f,A_f)}}{Z} \delta((\pi, A) - \Phi(\pi_f, A_f)) \min\{1, \frac{\mathrm{e}^{-H(\pi,A)}}{\mathrm{e}^{-H(\pi_f,A_f)}}\} \tag{8.12}$$

Now, since the map is symplectic, $d\pi_f dA_f = d\pi dA$. Since the map is also reversible the following holds:

$$\begin{aligned} \delta((\pi_f, A_f) - \Phi(\pi, A)) &= \frac{1}{|\det D\Phi|} \delta(\Phi(\pi_f, A_f) - (\pi, A)) \qquad (8.13) \\ &= \delta((\pi, A) - \Phi(\pi_f, A_f)) \qquad (8.14) \end{aligned}$$

Let's now suppose that $\frac{e^{-H(\pi,A)}}{e^{-H(\pi_f,A_f)}} < 1$. Then

$$p(\pi_f, A_f)T_2((\pi_f, A_f) \to (\pi, A)) = d\pi_f dA_f \frac{e^{-H(\pi_f,A_f)}}{Z}\delta((\pi, A) - \Phi(\pi_f, A_f))\min\{1, \frac{e^{-H(\pi,A)}}{e^{-H(\pi_f,A_f)}}\} \tag{8.15}$$

$$\begin{aligned} &= d\pi dA \frac{e^{-H(\pi_f,A_f)}}{Z}\delta((\pi, A) - \Phi(\pi_f, A_f))\frac{e^{-H(\pi,A)}}{e^{-H(\pi_f,A_f)}} \\ &= d\pi dA \frac{e^{-H(\pi,A)}}{Z}\delta((\pi_f, A_f) - \Phi(\pi, A)) \\ &= d\pi dA \frac{e^{-H(\pi,A)}}{Z}\delta((\pi_f, A_f) - \Phi(\pi, A))\min\{1, \frac{e^{-H(\pi_f,A_f)}}{e^{-H(\pi,A)}}\} \\ &= p(\pi, A)T_2((\pi, A) \to (\pi_f, A_f)) \end{aligned}$$

It's clear that the same holds when $\frac{e^{-H(\pi,A)}}{e^{-H(\pi_f,A_f)}} > 1$ and so T_2 also satisfies detailed balance. Therefore our HMC algorithm satisfies detailed balance and therefore converges to the canonical distribution with respect to $H(\pi, A)$, which induces a sample of the distribution desired $p(A) = \frac{e^{-\text{Re } S_{eff}(A)}}{\int DA \ e^{-\text{Re } S_{eff}}}$ upon a marginalization over π.

8.2 Selecting Momenta and Symplectic Evolution

Now that we have demonstrated that the HMC satisfies detailed balance and therefore converges to the canonical distribution, we explain how momenta are selected and we discuss the symplectic, reversible map Φ. To select momenta according to the probability

$$\text{Prob}_1(\pi) = p(\pi)D\pi = \frac{e^{-\frac{1}{2}\pi^T(J^\dagger J)^{-1}\pi}D\pi}{\int D\pi \ e^{-\frac{1}{2}\pi^T(J^\dagger J)^{-1}\pi}} \tag{8.16}$$

it suffices to select a momentum η distributed as

$$\text{Prob}_2(\eta) = \frac{e^{-\frac{1}{2}\eta^T\eta}D\eta}{\int D\eta e^{-\frac{1}{2}\eta^T\eta}} \tag{8.17}$$

and then define $\eta = (J^\dagger J)^{-1/2}\pi$. Indeed, consider a particular momentum π^*. The probability that this momentum is selected is the probability that $\eta = (J^\dagger J)^{-1/2}\pi^*$. So

$$\begin{aligned}
\mathrm{Prob}(\pi^*) &= \mathrm{Prob}_2(\eta = (J^\dagger J)^{-1/2}\pi^*) \\
&= \frac{\mathrm{e}^{-\frac{1}{2}\pi^{*T}(J^\dagger J)^{-1}\pi^*} D\eta}{\int D\eta \mathrm{e}^{-\frac{1}{2}\pi^{*T}(J^\dagger J)^{-1}\pi^*}} \\
&= \frac{\mathrm{e}^{-\frac{1}{2}\pi^{*T}(J^\dagger J)^{-1}\pi^*} D\eta \det\left(J^\dagger J\right)^{1/2}}{\int D\eta \det(J^\dagger J)^{1/2} \mathrm{e}^{-\frac{1}{2}\pi^{*T}(J^\dagger J)^{-1}\pi^*}} \\
&= \frac{\mathrm{e}^{-\frac{1}{2}\pi^{*T}(J^\dagger J)^{-1}\pi^*} D\pi}{\int D\pi\ \mathrm{e}^{-\frac{1}{2}\pi^{*T}(J^\dagger J)^{-1}\pi^*}} \\
&= \mathrm{Prob}_1(\pi^*)
\end{aligned} \tag{8.18}$$

The benefit of this approach is that there are exist very efficient algorithms for sampling from Gaussian distributions, so the sampling of the η variables is trivial. Furthermore, due to the sparsity of the jacobian, the multiplication required to transform from η to $\pi = (J^\dagger J)^{1/2}\eta$ is cheap. In fact, the multiplication is implemented analtically:

$$\pi_{0x} = \frac{1}{1 + f'(A_{0x})^2}\eta_{0x} \tag{8.19}$$

$$\pi_{1x} = \eta_{1x} \tag{8.20}$$

$$\pi_{2x} = \eta_{2x} \tag{8.21}$$

$$\tag{8.22}$$

In our calculations, we use an symplectic map Φ which is implicit. Φ is the composition of the following two maps

$$\begin{bmatrix} A_{N+1/2} \\ \pi_{N+1/2} \end{bmatrix} = \begin{bmatrix} A_N + \frac{\Delta t}{2} \nabla_\pi H(\pi_N, A_{N+1/2}) \\ \pi_N - \frac{\Delta t}{2} \nabla_A H(\pi_N, A_{N+1/2}) \end{bmatrix} \tag{8.23}$$

$$\begin{bmatrix} A_{N+1} \\ \pi_{N+1} \end{bmatrix} = \begin{bmatrix} A_{N+1/2} + \frac{\Delta t}{2} \nabla_\pi H(\pi_{N+1}, A_{N+1/2}) \\ \pi_{N+1/2} - \frac{\Delta t}{2} \nabla_A H(\pi_{N+1}, A_{N+1/2}) \end{bmatrix} \tag{8.24}$$

The first step is typically called an "Euler A" integration while the second step is called an "Euler B" integration [115]. Here Δt is a parameter which is chosen at the start of the simulation and $\nabla_\pi H$ is a vector with components $\frac{\partial H}{\partial \pi_i}$ and $\nabla_A H$ is a vector with components $\frac{\partial H}{\partial A_i}$. These maps are discrete approximations to Hamilton's equations. Since they are discretized, they will not conserve the Hamiltonian exactly. In spite of this discretization, this map is *exactly* reversible and symplectic, and a consequence of the reversibility is that the error in the Hamiltonian at a given time step is $\mathcal{O}(\Delta t^2)$. This quadratic dependence on Δt leads to an approximate conservation of H and therefor a high acceptance rate. Furthermore, note that Eq. (8.23) is an implicit equation in $A_{N+1/2}$ while Eq. (8.24) is an implicit equation in π_{N+1}. Therefore for this map to be well defined, Δt must be small enough that the implicit equation is well defined on all of phase space. We find this to be easily satisfied for $\Delta t \sim \mathcal{O}(10^{-1})$ in our simulations.

We presently demonstrate that our map $\Phi = \text{Euler B} \circ \text{Euler A}$ is symplectic and time reversible. To show that Φ is symplectic, we use the techniques of differential forms described in detail in [115]. The theorem we make use of is the following

Theorem 1 *If $dq \wedge dp$ is the canoncial 2-form on phase space, acting on pairs of vectors ζ and η as $(dq \wedge dp)(\zeta, \eta) = \zeta^T j \eta = \zeta^T \begin{bmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix} \eta$ if and $\begin{bmatrix} q \\ p \end{bmatrix} \longrightarrow \begin{bmatrix} \hat{q} \\ \hat{p} \end{bmatrix}$*

under some map Φ*, then* Φ *is symplectic if* $dq \wedge dp = d\hat{q} \wedge d\hat{p}$ *[115].*

This theorem is useful because all that is required to check the symplecticness of a map is the computation of a few differentials. To show that our map (Eq. (8.23) and Eq. (8.24)) is symplectic, it suffices to check that each step is symplectic because the composition of two symplectic maps is symplectic. Computing the differential of Eq. (8.23)

$$\begin{bmatrix} dA_{N+1/2} \\ d\pi_{N+1/2} \end{bmatrix} = \begin{bmatrix} dA_N + \frac{\Delta t}{2}\Big[\nabla_\pi\nabla_\pi H(\pi_N, A_{N+1/2})d\pi_N + \nabla_\pi\nabla_A H(\pi_N, A_{N+1/2})dA_{N+1/2}\Big] \\ d\pi_N - \frac{\Delta t}{2}\Big[\nabla_\pi\nabla_A H(\pi_N, A_{N+1/2})d\pi_N + \nabla_A\nabla_A H(\pi_N, A_{N+1/2})dA_{N+1/2}\Big] \end{bmatrix} \tag{8.25}$$

Rearranging, one finds

$$dA_{N+1/2} = (1 - \frac{\Delta t}{2}\nabla_\pi\nabla_A H(\pi_N, A_{N+1/2}))^{-1}\Big[dA_N + \frac{\Delta t}{2}\nabla_\pi\nabla_\pi H(\pi_N, A_{N+1/2})d\pi_N\Big] \tag{8.26}$$

$$d\pi_{N+1/2} = d\pi_N(1 - \frac{\Delta t}{2}\nabla_\pi\nabla_A H(\pi_N, A_{N+1/2})) - \frac{\Delta t}{2}\nabla_A\nabla_A H(\pi_N, A_{N+1/2})dA_{N+1/2} \tag{8.27}$$

Using the fact that $dA_{N+1/2} \wedge dA_{N+1/2} = 0$ one finds that

$$\begin{aligned} dA_{N+1/2} \wedge d\pi_{N+1/2} &= dA_{N+1/2} \wedge \Big[d\pi_N(1 - \frac{\Delta t}{2}\nabla_\pi\nabla_A H(\pi_N, A_{N+1/2}))\Big] \qquad (8.28) \\ &= \Big[dA_{N+1/2}(1 - \frac{\Delta t}{2}\nabla_\pi\nabla_A H(\pi_N, A_{N+1/2}))\Big] \wedge d\pi_N \\ &= \Big[dA_N + \frac{\Delta t}{2}\nabla_\pi\nabla_\pi H(\pi_N, A_{N+1/2})d\pi_N\Big] \wedge d\pi_N \\ &= dA_N \wedge d\pi_N \end{aligned}$$

where in the last line we have used $d\pi_N \wedge d\pi_N = 0$. Therefore the first step in our discretized Hamiltonian evolution is symplectic. A nearly identical set of manipulations demonstrates that the second step is also symplectic.

The final ingredient needed to show that out numerical method converges to the canonical distribution is to demonstrate that Φ is time reversal symmetric. To do so it suffices to show that the method is self-adjoint [115]. If one labels a method as $\psi_{\Delta t}$, for example $\psi_{\Delta t}$ could be the Euler A method defined in Eq. (8.23), then the adjoint method is defined as

$$\psi^*_{\Delta t} \equiv \psi^{-1}_{-\Delta t} \tag{8.29}$$

If we can show that the adjoint of Euler A is Euler B, then the proof is finished because

$$(\text{Euler A} \circ \text{Euler B})^* = (\text{Euler B}^* \circ \text{Euler A}^*) = \text{Euler A} \circ \text{Euler B} \tag{8.30}$$

This is indeed the case. To begin, let $\psi_{\Delta t}$ be the Euler A method and note that any adjoint method satisfies $\psi_{-\Delta t}(\psi^*_{\Delta t}(A, \pi)) = (A, \pi)$. If we denote $\psi^*_{\Delta t}(A, \pi) = (A^*, \pi^*)$ then

$$\begin{bmatrix} A \\ \pi \end{bmatrix} = \psi_{-\Delta t}(A^*, \pi^*) = \begin{bmatrix} A^* - \Delta t \nabla_\pi H(\pi^*, \mathcal{A}) \\ \pi^* + \Delta t \nabla_A H(\pi^*, \mathcal{A}) \end{bmatrix} \tag{8.31}$$

where we have defined $\mathcal{A}$ to be the position components of $\psi_{-\Delta t}(A^*, \pi^*)$. We have then that $\mathcal{A} = A$, and so

$$\begin{bmatrix} A \\ \pi \end{bmatrix} = \begin{bmatrix} A^* - \Delta t \nabla_\pi H(\pi^*, A) \\ \pi^* + \Delta t \nabla_A H(\pi^*, A) \end{bmatrix} \tag{8.32}$$

Therefore

$$\psi^*_{\Delta t}(A, \pi) = \begin{bmatrix} A^* \\ \pi^* \end{bmatrix} = \begin{bmatrix} A + \Delta t \nabla_\pi H(\pi^*, A) \\ \pi - \Delta t \nabla_A H(\pi^*, A) \end{bmatrix} \tag{8.33}$$

This the Euler B method. Therefore Euler B is the adjoint of Euler A and so our discretized flow map is self adjoint and therefore time reversible.

8.3 Integrator Details

In this section we detail how we implement our Euler B $\circ$ Euler A integrator. To begin, here are the relevant derivatives used to compute the discretized time evolution:

$$\frac{\partial H}{\partial \pi_{0x}} = \frac{\pi_{0x}}{1+f'(A_{0x})^2}, \ \frac{\partial H}{\partial \pi_{1x}} = \pi_{1x}, \ \frac{\partial H}{\partial \pi_{2x}} = \pi_{2x} \tag{8.34}$$

and

$$\begin{aligned}
\frac{\partial H}{\partial A_{0x}} &= \big(-\pi_{0x}^2\big)\frac{f'(A_{0x})f''(A_{0x})}{(1+f'(A_{0x})^2)^2} + \mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{0x}}(\tilde{A}(A))(1+if'(A_{0x}))\Big) \\
\frac{\partial H}{\partial A_{1x}} &= \mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{1x}}(\tilde{A}(A))\Big) \\
\frac{\partial H}{\partial A_{2x}} &= \mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{2x}}(\tilde{A}(A))\Big)
\end{aligned} \tag{8.35}$$

The Hamiltonian evolution is not trivial to compute because $\partial S/\partial A_{\mu x}$ appears, which requires a calculation of an inverse dirac matrix which is an expensive operation. Our procedure for computing the discrete map is designed to minimize the number of times we comute $\partial S/\partial A_{\mu x}$. For Euler B $\circ$Euler A, the dirac matrix needs to be inverted only once, while for Euler A $\circ$Euler B it needs to be computed twice.

To perform the Euler A integration we begin with a phase space configuration (A^N, π^N) then evolve according to:

$$\begin{bmatrix} A_{N+1/2} \\ \pi_{N+1/2} \end{bmatrix} = \begin{bmatrix} A_N + \frac{\Delta t}{2}\nabla_\pi H(\pi_N, A_{N+1/2}) \\ \pi_N - \frac{\Delta t}{2}\nabla_A H(\pi_N, A_{N+1/2}) \end{bmatrix} \tag{8.36}$$

In this method, the $A^{N+1/2}$ is an implicit calculation while $\pi^{N+1/2}$ is explicit. We first compute $A_{N+1/2}$. This is trivial for the space-like components:

$$A_{k,x}^{N+1/2} = A_{kx}^{N} + \frac{\Delta t}{2}\pi_{kx}^{N} \tag{8.37}$$

as there is no implicit equation to solve. For the timelike components we need to solve

$$A_{0,x}^{N+1/2} = A_{0,x}^{N} + \frac{\Delta t}{2}\Big[\frac{\pi_{0,x}^{N}}{1+f'(A_{0,x}^{N+1/2})^{2}}\Big] \tag{8.38}$$

which *is* implicit. This is done by finding the zero of the function

$$G(A; A_{0,x}^{N}, \pi_{0x}^{N}) = A - A_{0,x}^{N} - \frac{\Delta t}{2}\Big[\frac{\pi_{0,x}^{N}}{1+f'(A)^{2}}\Big] \tag{8.39}$$

closest to A_{0x}^{N} using the iterative Newton Solver $x_{i+1} = x_i - \frac{G(x_i)}{dG/dx(x_i)}$. With $A_{0x}^{N+1/2}$ in hand, $\pi_{0x}^{N+1/2}$ is obtained by computing the explicit relations

$$\pi_{0x}^{N+1/2} = \pi_{0x}^{N} - \frac{\Delta t}{2}\Big[\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{0x}}(\tilde{A}(A^{N+1/2}))(1+if'(A_{0x}^{N+1/2}))\Big) - \big({\pi_{0x}^{N}}^{2}\big)\frac{f'(A_{0x}^{N+1/2})f''(A_{0x}^{N+1/2})}{(1+f'(A_{0x}^{N+1/2})^{2})^{2}}\Big] \tag{8.40}$$

$$\pi_{1x}^{N+1/2} = \pi_{1x}^{N} - \frac{\Delta t}{2}\Big[\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{1x}}(\tilde{A}(A^{N+1/2}))\Big)\Big]$$

$$\pi_{2x}^{N+1/2} = \pi_{2x}^{N} - \frac{\Delta t}{2}\Big[\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{2x}}(\tilde{A}(A^{N+1/2}))\Big)\Big]$$

These relations can be computed with a single calculation of the gradient of the action. The second step of the integration, Euler B, is

$$\begin{bmatrix} A_{N+1} \\ \pi_{N+1} \end{bmatrix} = \begin{bmatrix} A_{N+1/2} + \frac{\Delta t}{2}\nabla_{\pi}H(\pi_{N+1}, A_{N+1/2}) \\ \pi_{N+1/2} - \frac{\Delta t}{2}\nabla_{A}H(\pi_{N+1}, A_{N+1/2}) \end{bmatrix} \tag{8.41}$$

This time A^{N+1} is explicit while π^{N+1} is implicit. Once again, the space-like components are easy to compute

$$\begin{aligned} \pi_{1x}^{N+1} &= \pi_{1x}^{N+1/2} - \frac{\Delta t}{2}\Big[\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{1x}}(\tilde{A}(A^{N+1/2}))\Big)\Big] \\ \pi_{2x}^{N+1} &= \pi_{2x}^{N+1/2} - \frac{\Delta t}{2}\Big[\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{2x}}(\tilde{A}(A^{N+1/2}))\Big)\Big] \end{aligned} \tag{8.42}$$

and require no implicit solutions. Note that the derivative needed in Eq. (8.42) is the same as the derivative needed in Eq. (8.43) and therefore need not be computed twice. This gives the Monte Carlo a factor of 2 speed-up relative to the less efficient Euler A $\circ$ Euler B method, more commonly called the "Störmer-Verlet" method [115]. The updated momenta π_{0x}^{N+1} is obtained by solving

$$\pi_{0x}^{N+1} = \pi_{0x}^{N+1/2} - \frac{\Delta t}{2}\Big[\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{0x}}(\tilde{A}(A^{N+1/2}))(1+if'(A_{0x}^{N+1/2}))\Big) - \big(\pi_{0x}^{N+1^2}\big)\frac{f'(A_{0x}^{N+1/2})f''(A_{0x}^{N+1/2})}{(1+f'(A_{0x}^{N+1/2})^2)^2}\Big]$$

This is just a quadratic equation. If we denote

$$\begin{aligned} A &= \frac{\Delta t}{2}\Big[\frac{f'(A_{0x}^{N+1/2})f''(A_{0x}^{N+1/2})}{(1+f'(A_{0x}^{N+1/2})^2)^2}\Big] \\ B &= -1 \\ C &= \pi_{0x}^{N+1/2} - \frac{\Delta t}{2}\mathrm{Re}\,\Big(\frac{\partial S}{\partial A_{0x}}(\tilde{A}(A^{N+1/2}))(1+if'(A_{0x}^{N+1/2}))\Big) \end{aligned} \tag{8.43}$$

then $\pi_{0x}^{N+1} = \frac{1-\sqrt{1-4AC}}{2A}$ is the correct root. Finally, A^{N+1} is

$$\begin{aligned} A_{0,x}^{N+1} &= A_{0,x}^{N+1/2} + \frac{\Delta t}{2}\Big[\frac{\pi_{0,x}^{N+1}}{1+{f'(A_{0,x}^{N+1/2})}^2}\Big] \\ A_{1,x}^{N+1} &= A_{1x}^{N+1/2} + \frac{\Delta t}{2}\pi_{1x}^{N+1} \\ A_{2,x}^{N+1} &= A_{2x}^{N+1/2} + \frac{\Delta t}{2}\pi_{2x}^{N+1} \end{aligned} \tag{8.44}$$

Chapter 9

Bibliography

[1] P. W. Anderson. More is different. *Science*, 177(4047):393–396, 1972.

[2] H. J. Rothe. Lattice gauge theories: An Introduction. *World Sci. Lect. Notes Phys.*, 43:1–381, 1992. [World Sci. Lect. Notes Phys.82,1(2012)].

[3] I. Montvay and G. Munster. *Quantum fields on a lattice.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1997.

[4] Christof Gattringer and Christian B. Lang. Quantum chromodynamics on the lattice. *Lect. Notes Phys.*, 788:1–343, 2010.

[5] Gordon Baym and Christopher Pethick. Physics of neutron stars. *Annual Review of Astronomy and Astrophysics*, 17(1):415–443, 1979.

[6] Gordon Baym, Tetsuo Hatsuda, Toru Kojo, Philip D. Powell, Yifan Song, and Tatsuyuki Takatsuka. From hadrons to quarks in neutron stars: a review. *Rept. Prog. Phys.*, 81(5):056902, 2018.

[7] Adam Burrows and James M. Lattimer. The birth of neutron stars. *Astrophys. J.*, 307:178–196, 1986.

[8] Mehran Kardar. *Statistical Physics of Fields.* Cambridge University Press, 2007.

[9] L.D. LANDAU and E.M. LIFSHITZ. Chapter i - the fundamental principles of statistical physics. In L.D. LANDAU and E.M. LIFSHITZ, editors, *Statistical Physics (Third Edition)*, pages 1 – 33. Butterworth-Heinemann, Oxford, third edition edition, 1980.

[10] Jan Smit. *Introduction to Quantum Fields on a Lattice*. Cambridge Lecture Notes in Physics. Cambridge University Press, 2002.

[11] Richard R. Silbar and Sanjay Reddy. Neutron stars for undergraduates. *American Journal of Physics*, 72(7):892–905, 2004.

[12] Charles J. Geyer. 1 introduction to markov chain monte carlo. 2011.

[13] Philippe de Forcrand. Simulating QCD at finite density. *PoS*, LAT2009:010, 2009.

[14] K. Splittorff and J. J. M. Verbaarschot. Approach to the thermodynamic limit in lattice qcd at $\mu \neq 0$. *Phys. Rev. D*, 77:014514, Jan 2008.

[15] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar. Numerical study of the two-dimensional hubbard model. *Phys. Rev. B*, 40:506–516, Jul 1989.

[16] Gert Aarts. Introductory lectures on lattice QCD at nonzero baryon number. *J. Phys. Conf. Ser.*, 706(2):022004, 2016.

[17] Matthias Troyer and Uwe-Jens Wiese. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.*, 94:170201, 2005.

[18] Andrei Alexandru, Gokce Basar, Paulo F. Bedaque, Gregory W. Ridgway, and Neill C. Warrington. Sign problem and Monte Carlo calculations beyond Lefschetz thimbles. *JHEP*, 05:053, 2016.

[19] Andrei Alexandru, Gokce Basar, Paulo F. Bedaque, Sohan Vartak, and Neill C. Warrington. Monte Carlo Study of Real Time Dynamics on the Lattice. *Phys. Rev. Lett.*, 117(8):081602, 2016.

[20] Andrei Alexandru, Gokce Basar, Paulo Bedaque, Gregory W. Ridgway, and Neill C. Warrington. Study of symmetry breaking in a relativistic Bose gas using the contraction algorithm. *Phys. Rev.*, D94(4):045017, 2016.

[21] Andrei Alexandru, Gokce Basar, Paulo F. Bedaque, Gregory W. Ridgway, and Neill C. Warrington. Monte Carlo calculations of the finite density Thirring model. *Phys. Rev.*, D95(1):014502, 2017.

[22] Andrei Alexandru, Paulo F. Bedaque, Henry Lamm, Scott Lawrence, and Neill C. Warrington. Fermions at Finite Density in 2+1 Dimensions with Sign-Optimized Manifolds. *Phys. Rev. Lett.*, 121(19):191602, 2018.

[23] Andrei Alexandru, Gokce Basar, Paulo F. Bedaque, and Gregory W. Ridgway. Schwinger-Keldysh formalism on the lattice: A faster algorithm and its application to field theory. *Phys. Rev.*, D95(11):114501, 2017.

[24] Edward Witten. Analytic Continuation Of Chern-Simons Theory. *AMS/IP Stud. Adv. Math.*, 50:347–446, 2011.

[25] Edward Witten. A New Look At The Path Integral Of Quantum Mechanics. 2010.

[26] J. Milnor, M. SPIVAK, and R. WELLS. *Morse Theory. (AM-51), Volume 51.* Princeton University Press, 1969.

[27] Liviu Nicolaescu. An invitation to morse theory. 01 2011.

[28] Frédéric Pham. Vanishing homologies and the n variable saddlepoint method. In *Singularities, Part 2 (Arcata, Calif., 1981)*, volume 40 of *Proc.*

Sympos. Pure Math., pages 319–333. Amer. Math. Soc., Providence, RI, 1983.

[29] Marco Cristoforetti, Francesco Di Renzo, and Luigi Scorzato. New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble. *Phys. Rev.*, D86:074506, 2012.

[30] Marco Cristoforetti, Luigi Scorzato, and Francesco Di Renzo. The sign problem and the lefschetz thimble. *Journal of Physics: Conference Series*, 432:012025, apr 2013.

[31] Marco Cristoforetti, Francesco Di Renzo, Abhishek Mukherjee, and Luigi Scorzato. Monte carlo simulations on the lefschetz thimble: Taming the sign problem. *Phys. Rev. D*, 88:051501, Sep 2013.

[32] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano. Hybrid monte carlo on lefschetz thimbles — a study of the residual sign problem. *Journal of High Energy Physics*, 2013(10):147, Oct 2013.

[33] Abhishek Mukherjee, Marco Cristoforetti, and Luigi Scorzato. Metropolis monte carlo integration on the lefschetz thimble: Application to a one-plaquette model. *Phys. Rev. D*, 88:051502, Sep 2013.

[34] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato. Quantum field theories on the Lefschetz thimble. *PoS*, LATTICE2013:197, 2014.

[35] M. Cristoforetti, F. Di Renzo, G. Eruzzi, A. Mukherjee, C. Schmidt, L. Scorzato, and C. Torrero. An efficient method to compute the residual phase on a lefschetz thimble. *Phys. Rev. D*, 89:114505, Jun 2014.

[36] Gert Aarts, Lorenzo Bongiovanni, Erhard Seiler, and Dénes Sexty. Some remarks on lefschetz thimbles and complex langevin dynamics. *Journal of High Energy Physics*, 2014(10):159, Oct 2014.

[37] Yuya Tanizaki. Lefschetz-thimble techniques for path integral of zero-dimensional $o(n)$ sigma models. *Phys. Rev. D*, 91:036002, Feb 2015.

[38] Yuya Tanizaki, Yoshimasa Hidaka, and Tomoya Hayata. Lefschetz-thimble analysis of the sign problem in one-site fermion model. *New J. Phys.*, 18(3):033002, 2016.

[39] Kenji Fukushima and Yuya Tanizaki. Hamilton dynamics for Lefschetz-thimble integration akin to the complex Langevin method. *PTEP*, 2015(11):111A01, 2015.

[40] Francesco Di Renzo and Giovanni Eruzzi. Thimble regularization at work: from toy models to chiral random matrix theories. *Phys. Rev.*, D92(8):085030, 2015.

[41] Kevin Zambello and Francesco Di Renzo. Towards Lefschetz thimbles regularization of heavy-dense QCD. In *36th International Symposium on Lattice Field Theory (Lattice 2018) East Lansing, MI, United States, July 22-28, 2018*, 2018.

[42] Francesco Di Renzo. Simulating lattice field theories on multiple thimbles. *EPJ Web Conf.*, 175:07015, 2018.

[43] F. Di Renzo and G. Eruzzi. One-dimensional qcd in thimble regularization. *Phys. Rev. D*, 97:014503, Jan 2018.

[44] Giovanni Eruzzi and Francesco Di Renzo. Thimble regularization at work besides toy models: from Random Matrix Theory to Gauge Theories. *PoS*, LATTICE2015:188, 2016.

[45] Francesco Di Renzo and Giovanni Eruzzi. Thimble regularization at work for Gauge Theories: from toy models onwards. *PoS*, LATTICE2015:189, 2016.

[46] Stefan Bluecher, Jan M. Pawlowski, Manuel Scherzer, Mike Schlosser, Ion-Olimpiu Stamatescu, Sebastian Syrkowski, and Felix P. G. Ziegler. Reweighting Lefschetz Thimbles. *SciPost Phys.*, 5(5):044, 2018.

[47] Christian Schmidt and Felix Ziesché. Simulating low dimensional QCD with Lefschetz thimbles. *PoS*, LATTICE2016:076, 2017.

[48] Andrei Alexandru, Paulo F. Bedaque, Henry Lamm, and Scott Lawrence. Finite-Density Monte Carlo Calculations on Sign-Optimized Manifolds. *Phys. Rev.*, D97(9):094510, 2018.

[49] G. Parisi. ON COMPLEX PROBABILITIES. *Phys. Lett.*, 131B:393–395, 1983.

[50] Gert Aarts and Ion-Olimpiu Stamatescu. Stochastic quantization at finite chemical potential. *Journal of High Energy Physics*, 2008(09):018–018, sep 2008.

[51] Gert Aarts. Can stochastic quantization evade the sign problem? the relativistic bose gas at finite chemical potential. *Phys. Rev. Lett.*, 102:131601, Apr 2009.

[52] Gert Aarts and Frank A. James. On the convergence of complex langevin dynamics: the three-dimensional xy model at finite chemical potential. *Journal of High Energy Physics*, 2010(8):20, Aug 2010.

[53] Gert Aarts, Frank A. James, Erhard Seiler, and Ion-Olimpiu Stamatescu. Complex langevin: etiology and diagnostics of its main problem. *The European Physical Journal C*, 71(10):1756, Oct 2011.

[54] J. B. Kogut and D. K. Sinclair. Applying Complex Langevin Simulations to Lattice QCD at Finite Density. 2019.

[55] Keitaro Nagata, Jun Nishimura, and Shinji Shimasaki. Complex langevin calculations in finite density qcd at large μ/t with the deformation technique. *Phys. Rev. D*, 98:114513, Dec 2018.

[56] DÃ©nes Sexty. Simulating full qcd at nonzero density using the complex langevin equation. *Physics Letters B*, 729:108 – 111, 2014.

[57] Bloch, Jacques, Glesaaen, Jonas, Philipsen, Owe, Verbaarschot, Jacobus, and Zafeiropoulos, Savvas. Complex langevin simulations of a finite density matrix model for qcd. *EPJ Web Conf.*, 137:07030, 2017.

[58] J. E. Drut. Advances in non-relativistic matter via complex Langevin approaches. *J. Phys. Conf. Ser.*, 1041(1):012005, 2018.

[59] Jun Nishimura and Asato Tsuchiya. Complex Langevin analysis of the space-time structure in the Lorentzian type IIB matrix model. 2019.

[60] Gert Aarts, Kirill Boguslavski, Manuel Scherzer, Erhard Seiler, Dénes Sexty, and Ion-Olimpiu Stamatescu. Getting even with CLE. *EPJ Web Conf.*, 175:14007, 2018.

[61] Jacques Bloch, Jonas Glesaaen, Jacobus Verbaarschot, and Savvas Zafeiropoulos. Progress on Complex Langevin simulations of a finite density matrix model for QCD. *EPJ Web Conf.*, 175:07034, 2018.

[62] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. Toward solving the sign problem with path optimization method. *Phys. Rev. D*, 96:111501, Dec 2017.

[63] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. Application of a neural network to the sign problem via the path optimization method. *PTEP*, 2018(2):023B04, 2018.

[64] Kouji Kashiwa, Yuto Mori, and Akira Ohnishi. Controlling the model sign problem via the path optimization method: Monte carlo approach to a qcd effective model with polyakov loop. *Phys. Rev. D*, 99:014033, Jan 2019.

[65] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. Path optimization in 0+1 dimensional QCD at finite density. 2019.

[66] Christof Gattringer. Flux representation of an effective Polyakov loop model for QCD thermodynamics. *Nucl. Phys.*, B850:242–252, 2011.

[67] Ydalia Delgado, Hans Gerd Evertz, and Christof Gattringer. Solving the Complex Phase Problem in a QCD Related Model. *Acta Phys. Polon. Supp.*, 4:703–708, 2011.

[68] Ydalia Delgado Mercado, Hans Gerd Evertz, and Christof Gattringer. Worm algorithms for the 3-state Potts model with magnetic field and chemical potential. *Comput. Phys. Commun.*, 183:1920–1927, 2012.

[69] Ydalia Delgado Mercado and Christof Gattringer. Monte Carlo simulation of the SU(3) spin model with chemical potential in a flux representation. *Nucl. Phys.*, B862:737–750, 2012.

[70] Christof Gattringer and Thomas Kloiber. Lattice study of the Silver Blaze phenomenon for a charged scalar ϕ^4 field. *Nucl. Phys.*, B869:56–73, 2013.

[71] Christof Gattringer and Alexander Schmidt. Gauge and matter fields as surfaces and loops - an exploratory lattice study of the Z(3) Gauge-Higgs model. *Phys. Rev.*, D86:094506, 2012.

[72] Ydalia Delgado Mercado, Christof Gattringer, and Alexander Schmidt. Dual Lattice Simulation of the Abelian Gauge-Higgs Model at Finite Density: An Exploratory Proof of Concept Study. *Phys. Rev. Lett.*, 111(14):141601, 2013.

[73] Thomas Kloiber and Christof Gattringer. Scalar QED$_2$ with a topological term – a lattice study in a dual representation. *PoS*, LATTICE2014:345, 2014.

[74] Christof Gattringer, Thomas Kloiber, and Vasily Sazonov. Solving the sign problems of the massless lattice Schwinger model with a dual formulation. *Nucl. Phys.*, B897:732–748, 2015.

[75] Falk Bruckmann, Christof Gattringer, Thomas Kloiber, and Tin Sulejmanpasic. Dual lattice representations for O(N) and CP(N1) models with a chemical potential. *Phys. Lett.*, B749:495–501, 2015. [Erratum: Phys. Lett.B751,595(2015)].

[76] Christof Gattringer, Daniel Göschl, and Tin Sulejmanpasic. Dual simulation of the 2d U(1) gauge Higgs model at topological angle $\theta = \pi$: Critical endpoint behavior. *Nucl. Phys.*, B935:344–364, 2018.

[77] Shailesh Chandrasekharan and Uwe-Jens Wiese. Meron cluster solution of a fermion sign problem. *Phys. Rev. Lett.*, 83:3116–3119, 1999.

[78] W. Bietenholz, A. Pochinsky, and U. J. Wiese. Meron-cluster simulation of the θ vacuum in the 2d o(3) model. *Phys. Rev. Lett.*, 75:4524–4527, Dec 1995.

[79] S. Chandrasekharan, J. Cox, K. Holland, and U. J. Wiese. Meron cluster simulation of a chiral phase transition with staggered fermions. *Nucl. Phys.*, B576:481–500, 2000.

[80] S. Chandrasekharan, B. Scarlet, and U. J. Wiese. Meron cluster simulation of quantum spin ladders in a magnetic field. 1999.

[81] Shailesh Chandrasekharan. A Chiral phase transition using a fermion cluster algorithm. *Chin. J. Phys.*, 38:696–706, 2000.

[82] J. Cox, C. Gattringer, K. Holland, B. Scarlet, and U.-J. Wiese. Meron-cluster solution of fermion and other sign problems. *Nuclear Physics B - Proceedings Supplements*, 83-84:777 – 791, 2000. Proceedings of the XVIIth International Symposium on Lattice Field Theory.

[83] Mark G. Alford, S. Chandrasekharan, J. Cox, and U. J. Wiese. Solution of the complex action problem in the Potts model for dense QCD. *Nucl. Phys.*, B602:61–86, 2001.

[84] Shailesh Chandrasekharan and James C. Osborn. Kosterlitz-Thouless universality in a fermionic system. *Phys. Rev.*, B66:045113, 2002.

[85] Shailesh Chandrasekharan. Fermion cluster algorithms. *Nuclear Physics B - Proceedings Supplements*, 83-84:774 – 776, 2000. Proceedings of the XVIIth International Symposium on Lattice Field Theory.

[86] Shailesh Chandrasekharan and Anyi Li. Fermion bag approach to the sign problem in strongly coupled lattice QED with Wilson fermions. *JHEP*, 01:018, 2011.

[87] Shailesh Chandrasekharan and Anyi Li. Fermion bags, duality and the three dimensional massless lattice Thirring model. *Phys. Rev. Lett.*, 108:140404, 2012.

[88] Shailesh Chandrasekharan and Anyi Li. Fermion bag solutions to some sign problems in four-fermion field theories. *Phys. Rev.*, D85:091502, 2012.

[89] Shailesh Chandrasekharan and Anyi Li. Quantum critical behavior in three dimensional lattice Gross-Neveu models. *Phys. Rev.*, D88:021701, 2013.

[90] Emilie Fulton Huffman and Shailesh Chandrasekharan. Solution to sign problems in half-filled spin-polarized electronic systems. *Phys. Rev.*, B89(11):111101, 2014.

[91] Venkitesh Ayyar, Shailesh Chandrasekharan, and Jarno Rantaharju. Benchmark results in the 2D lattice Thirring model with a chemical potential. *Phys. Rev.*, D97(5):054501, 2018.

[92] Andrei Alexandru, Gök çe Başar, Paulo F. Bedaque, and Gregory Ridgway. Schwinger-keldysh formalism on the lattice: A faster algorithm and its application to field theory. *Phys. Rev. D*, 95:114501, Jun 2017.

[93] Andrei Alexandru, Paulo F. Bedaque, and Neill C. Warrington. Spin polarized nonrelativistic fermions in $1 + 1$ dimensions. *Phys. Rev. D*, 98:054514, Sep 2018.

[94] Andrei Alexandru, Gök çe Başar, Paulo F. Bedaque, Gregory W. Ridgway, and Neill C. Warrington. Fast estimator of jacobians in the monte carlo integration on lefschetz thimbles. *Phys. Rev. D*, 93:094514, May 2016.

[95] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[96] Walter E. Thirring. A Soluble relativistic field theory? *Annals Phys.*, 3:91–112, 1958. [,509(1958)].

[97] C. R. Hagen. New solutions of the thirring model. *Il Nuovo Cimento B (1965-1970)*, 51(1):169–186, Sep 1967.

[98] V. Glaser. An explicit solution of the thirring model. *Il Nuovo Cimento (1955-1965)*, 9(6):990–1006, Sep 1958.

[99] V. E. Korepin. Direct calculation of the s matrix in the massive thirring model. *Theoretical and Mathematical Physics*, 41(2):953–967, Nov 1979.

[100] Sidney Coleman. Quantum sine-gordon equation as the massive thirring model. *Phys. Rev. D*, 11:2088–2097, Apr 1975.

[101] E. Shuster and D.T. Son. On finite-density qcd at large nc. *Nuclear Physics B*, 573(1):434 – 446, 2000.

[102] Ian Stewart and David Tall. *Complex Analysis.* Cambridge University Press, 2 edition, 2018.

[103] J. R. Cash and Alan H. Karp. A variable order runge-kutta method for initial value problems with rapidly varying right-hand sides. *ACM Trans. Math. Softw.*, 16(3):201–222, September 1990.

[104] Luca Guido Molinari. Determinants of block tridiagonal matrices. *Linear Algebra and its Applications*, 429(8):2221 – 2226, 2008.

[105] Jan M. Pawlowski, Ion-Olimpiu Stamatescu, and Christian Zielinski. Simple QED- and QCD-like Models at Finite Density. *Phys. Rev.*, D92(1):014508, 2015.

[106] *NIST Digital Library of Mathematical Functions.* http://dlmf.nist.gov/, Release 1.0.22 of 2019-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[107] David B. Kaplan. Chiral Symmetry and Lattice Fermions. In *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Proceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009*, pages 223–272, 2009.

[108] Thomas D. Cohen. Functional integrals for qcd at nonzero chemical potential and zero density. *Phys. Rev. Lett.*, 91:222001, Nov 2003.

[109] Yuya Tanizaki, Yoshimasa Hidaka, and Tomoya Hayata. Lefschetz-thimble approach to the Silver Blaze problem of one-site fermion model. *PoS*, LATTICE2016:030, 2016.

[110] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

[111] Jan M. Pawlowski and Christian Zielinski. Thirring model at finite density in 2+1 dimensions with stochastic quantization. *Phys. Rev.*, D87(9):094509, 2013.

[112] Holger Gies and Lukas Janssen. Uv fixed-point structure of the three-dimensional thirring model. *Phys. Rev. D*, 82:085018, Oct 2010.

[113] Simon Duane, A.D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216 – 222, 1987.

[114] Radford M. Neal. MCMC using Hamiltonian dynamics. *arXiv e-prints*, page arXiv:1206.1901, Jun 2012.

[115] Benedict J Leimkuhler and Sebastian Reich. *Simulating Hamiltonian dynamics*. Cambridge monographs on applied and computational mathematics. Cambridge Univ., Cambridge, 2004.